

Delaval, Jan

91408

From: Ewoldt, Gerald
Sent: Tuesday, April 15, 2003 10:09 AM
To: Delaval, Jan
Subject: seq search for 09/521,527

Jan,
Please search SEQ ID NO:1, closed and including interference.

Thanks,
Gerry

G.R. Ewoldt, Ph.D.
Patent Examiner
CM1, 9B13 or 9E12
703-308-9805

Jan Delaval
Reference Librarian
Biotechnology & Chemical Library
CM1 1E07 -- 703-308-4483
jan.delaval@uspto.gov

GenCore version 5.1.4_P5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 15, 2003, 10:13:59 ; Search time 34 Seconds

(without alignments)
47.030 Million cell updates/sec

Title: US-09-521-527C-1

Perfect score: 60

Sequence: 1 AAEGLDTRFSG 12

Scoring table:

BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 204235

Minimum DB seq length: 0
Maximum DB seq length: 12

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A.Geneseq_101002:*

1: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*

2: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*

3: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*

4: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*

5: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*

6: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*

7: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*

8: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*

9: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*

10: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*

11: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*

12: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*

13: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*

14: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*

15: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*

16: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*

17: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*

18: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*

19: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*

20: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*

21: /SID52/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*

22: /SID52/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*

23: /SID52/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	60	100.0	12	21	AA824101	Human CD8 antigeni
2	54	90.0	11	21	AA824102	Human CD8 antigeni
3	40	66.7	8	17	AA89502	CD8 alpha chain re
4	40	66.7	8	21	AAV43646	Fragment of a huma
5	40	66.7	10	17	AA89509	CD8 antagonist #4.
6	40	66.7	10	21	AAV43649	Amino acid sequenc
7	32	53.3	11	22	AAW97622	Human peptide #897
8	26	43.3	9	23	ABG60421	Selective targetin
9	26	43.3	11	23	ABB78736	Chicken calstynteni
10	25	41.7	9	17	AAW49400	Human leucocyte an

11	25	41.7	11	18	AAW15769	Anion exchange sta
12	25	41.7	11	22	AAU03264	Fruit fly G protei
13	25	41.7	11	22	AAU03828	G protein-coupled
14	25	41.7	12	16	AA86812	Syndecan heparan-a
15	24	40.0	9	15	AA861860	MBP peptide 108, p
16	24	40.0	9	17	AAW49399	Human leucocyte an
17	24	40.0	9	22	AAW96631	Human peptide #190
18	24	40.0	9	23	ABB97050	Human tumour anti
19	24	40.0	10	18	AAW16461	Phosphopantetheiny
20	24	40.0	11	15	AA861960	MBP peptide 108, p
21	24	40.0	11	16	AA870696	Mesquite allergeni
22	24	40.0	11	21	AAV70378	Maize ZmKCS1 pepi
23	23	38.3	8	21	AA861861	Cadherin-8 cell ad
24	23	38.3	8	22	AA861861	Hepatoma-diagnosti
25	23	38.3	8	22	AA861861	Hepatoma-diagnosti
26	23	38.3	9	19	AAW63555	Beta (1 -> 4)-N-ac
27	23	38.3	9	20	AAV55363	HLA binding plu-1
28	23	38.3	9	22	AA898572	Human TADG-15 pept
29	23	38.3	10	19	AAW57613	T-cell receptor CD
30	23	38.3	10	21	AAV86603	T-cell receptor co
31	23	38.3	10	22	AA895622	Human complementar
32	23	38.3	10	22	AA895624	Human complementar
33	23	38.3	10	22	AAU03247	Fruit fly G protei
34	23	38.3	10	22	AAU03267	Fruit fly G protei
35	23	38.3	10	22	AAU03844	G protein-coupled
36	23	38.3	10	23	AAU80241	Unique peptide der
37	23	38.3	10	23	AA897051	Human tumour anti
38	23	38.3	11	23	AA861674	Human KRP1 tryptic
39	22	36.7	6	21	AA820953	Bovine beta-lactog
40	22	36.7	6	22	AA862887	Mature bovine beta
41	22	36.7	6	23	AA862887	Peptide T7 derived
42	22	36.7	8	22	AAE12614	Murine Ob receptor
43	22	36.7	8	23	AAE23864	Mouse Ob receptor
44	22	36.7	9	22	AAE05202	Human HLA-A1 bindi
45	22	36.7	9	22	AAE05220	Human HLA-A3 bindi

ALIGNMENTS

RESULT 1	AA824101	standard; peptide: 12 AA.
ID	AA824101	
XX	AA824101	
AC	AA824101	
XX	AA824101	
XX	23-JAN-2001	(First entry)
DT	XX	
XX	XX	
DE	Human CD8 antigenic peptide.	
XX	XX	
KW	Human: CD8; antibody; antigenic; viral infection; tumour;	
KW	CD8+ cell; peripheral mononuclear blood cell.	
XX	XX	
OS	Homo sapiens.	
XX	XX	
PN	MO200055305-A1.	
XX	XX	
PD	21-SEP-2000.	
XX	XX	
PF	08-MAR-2000; 2000MO-US05898.	
XX	XX	
PR	12-MAR-1999; 99US-0124253.	
XX	XX	
PA	(ORTH) ORTHO-MCNETL PHARM INC.	
XX	XX	
PI	Leturcq D;	
XX	XX	
DR	WPI; 2000-618917/59.	
XX	XX	
PT	Isolating CD8+ cells from peripheral mononuclear blood cells for use as	
PT	vehicles for combating viral infections and tumors, comprises using a	
XX	CD8 specific antibody that does not activate the cells	

PS Claim 1: Page 15; 21pp; English.

CC The present invention describes a method for isolating CD8+ cells (1)

CC from peripheral mononuclear blood cells (PBMC) using an antibody which

CC specifically binds to the sequence AAEGLDTRFSG (1) or a portion of (1)

CC on CD8 molecules present on the surface of CD8+ cells but does not

CC activate the CD8+ cells once bound. Also described are: (1) a hybridoma

CC cell line (II) which produces a monoclonal antibody that specifically

CC binds to CD8 molecules present on the surface of CD8+ cells but does not

CC activate the CD8+ cells; (2) a monoclonal antibody (III) produced by

CC (III) (3) a polypeptide comprising (1), useful for generating (III);

CC (4) a population of CD8+ cells isolated by (1); and (5) a kit, comprising

CC (1) and an agent which causes the dissociation of CD8+ cell-antibody

CC complex. (1) is useful for isolating human CD8+ cells from a sample of

CC isolated peripheral mononuclear blood cells. The isolated CD8+ cells have

CC importance as vehicles for combating viral infections and tumours.

CC

SO Sequence 12 AA:

Query Match 100.0%; Score 60; DB 21; Length 12;

Best Local Similarity 100.0%; Pred. No. 1,4e-05;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAEGLDTRFSG 12

DB 1 AAEGLDTRFSG 12

RESULT 2

ID AAB24102 standard; peptide; 11 AA.

XX AAB24102;

AC AAB24102;

XX

DT 23-JAN-2001 (first entry)

XX

DE Human CD8 antigenic peptide CD8-3.

XX

KM Human: CD8; antibody; antigenic; viral infection; tumour;

KW CD8+ cell; peripheral mononuclear blood cell.

XX

OS Homo sapiens.

XX

PN WO200055305-A1.

XX

PD 21-SEP-2000.

XX

PF 08-MAR-2000; 2000WO-US05898.

XX

PR 12-MAR-1999; 99US-0124253.

XX

PA (ORTH) ORTHO-MCNEIL PHARM INC.

PI Leturcq D;

XX

DR WPI: 2000-618917/59.

XX

PT Isolating CD8+ cells from peripheral mononuclear blood cells for use as

PT vehicles for combating viral infections and tumors, comprises using a

PT CD8 specific antibody that does not activate the cells

XX

PS Disclosure: Page 8; 21pp; English.

XX

CC The present invention describes a method for isolating CD8+ cells (1)

CC from peripheral mononuclear blood cells (PBMC) using an antibody which

CC specifically binds to the sequence AAEGLDTRFSG (1) or a portion of (1)

CC on CD8 molecules present on the surface of CD8+ cells but does not

CC activate the CD8+ cells once bound. Also described are: (1) a hybridoma

CC cell line (II) which produces a monoclonal antibody that specifically

CC binds to CD8 molecules present on the surface of CD8+ cells but does not

CC activate the CD8+ cells; (2) a monoclonal antibody (III) produced by

CC (III) (3) a polypeptide comprising (1), useful for generating (III);

CC (4) a population of CD8+ cells isolated by (1); and (5) a kit, comprising

CC (III) and an agent which causes the dissociation of CD8+ cell-antibody

CC complex. (1) is useful for isolating human CD8+ cells from a sample of

CC isolated peripheral mononuclear blood cells. The isolated CD8+ cells have

CC importance as vehicles for combating viral infections and tumors. The

CC present sequence represents a human CD8 antigenic peptide designated

CC CD8-3, which is given in the exemplification of the present invention.

CC

SO Sequence 11 AA:

Query Match 90.0%; Score 54; DB 21; Length 11;

Best Local Similarity 100.0%; Pred. No. 0.00022;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAEGLDTRFS 11

DB 1 AAEGLDTRFS 11

RESULT 3

ID AAR98502 standard; peptide; 8 AA.

XX AAR98502;

AC AAR98502;

XX

DT 04-MAR-1997 (first entry)

XX

DE CD8 alpha chain residues 60-67.

XX

DE CD8 antagonist; inhibitor; T-cell activation; human; CD8 alpha chain;

KW thymic differentiation; transplantation; bone marrow; liver; heart; lung;

KW kidney; cornea; skin graft; graft versus host disease; therapy.

XX

OS Homo sapiens.

XX

PN WO9622106-A1.

XX

PD 25-JUL-1996.

XX

PF 17-JAN-1996; 96WO-US00310.

XX

PR 17-JAN-1995; 95US-0372952.

XX

PA (UYJE-) UNIV JEFFERSON THOMAS.

XX

PI Choksi S, Huang Z, Jameson BA, Kornigold R;

XX

DR WPI: 1996-354307/35.

XX

PT CD8 antagonist peptide(s) - used for inhibiting T cell activation,

PT partic. for treating transplant rejection or graft versus host

PT disease

XX

PS Claim 2: Page 28; 44pp; English.

XX

CC AAR98501-R98505 represent components of a CD8 antagonist of the

CC invention. These sequences are fragments of the human CD8 alpha chain

CC molecule (see AAR98514). This sequence corresponds to residues 60-67 of

CC the human CD8 sequence. CD8 plays a major role in the activation of

CC mature T-cells, and in the thymic differentiation process that leads to

CC CD8 expression. CD8 is expressed either as a homodimer (containing two

CC alpha chains) or as a heterodimer (an alpha and a beta chain). The CD8

CC CD82-like region is involved in regulating T-cell activation. The

CC antagonists of the invention comprise a molecular surface similar to at

CC least a portion of human CD8 molecular surface around the site of one of

CC these peptides. The compounds of the invention compete with CD8 so as to

CC inhibit T-cell activation. They can also be used to treat an individual

CC who is about to undergo, is undergoing, or has undergone a

CC transplantation procedure such as bone marrow, liver, heart, kidney,

CC lung, islets, or cornea transplantation, or skin grafts. The compounds

CC can also be used to treat an individual suspected of suffering from, or

CC susceptible to graft versus host disease.

XX

SO Sequence 8 AA:

Query Match 66.7%; Score 40; DB 17; Length 8;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AEGDITOR 9
 | | | | | | | |
 DB 1 AEGDITOR 8

RESULT 4
 AAY43646
 ID AAY43646 standard; peptide: 8 AA.

XX AAY43646;
 XX
 DT 11-FEB-2000 (first entry)

DE Fragment of a human CD8 antagonist peptide.

XX Extracellular domain; alpha chain; human CD8; T cell activation;
 KW cytotoxic T lymphocyte activation; CTL activation;

KM graft versus host disease; CD8 antagonist; transplant;

KW allogeneic bone marrow transplant.

OS Homo sapiens.

XX WO954345-A1.

XX 28-OCT-1999.

XX 21-APR-1999; 99WO-US08814.

XX 21-APR-1998; 98US-0082436.

XX (UYJE-) UNIV JEFFERSON THOMAS.

XX Korgold R, Huang Z, Choksi S;

DR WPI; 2000-013224/01.

PT Novel CD8 antagonists used to inhibit cytotoxic T lymphocytes and
 PT prevent immune responses -

PS Disclosure; Page 48; 67pp; English.

XX The present sequence represents a fragment of a human CD8 antagonist
 CC peptide. The specification describes peptides that are CD8 antagonists,
 CC and that compete with regions within the extracellular domain of the
 CC alpha chain of CD8 regions in intermolecular interactions that involve
 CC CD8 which are associated with cytotoxic T lymphocyte (CTL) activation.
 CC These peptides comprise a molecular surface that corresponds to a
 CC molecular surface of human CD8 at amino acids 73-76, 38-46, 53-56,
 CC 60-67 or 53-67. The peptides can be used in methods and compositions
 CC for therapeutically or prophylactically treating graft versus host
 CC disease in patients who have or about to undergo transplants, e.g.
 CC allogeneic bone marrow transplants. The methods and compositions are
 CC also used to treat individuals who are undergoing and/or who have
 CC undergone transplants, e.g. transplants of heart, liver, lung, islets
 CC or cornea, and skin grafts. The methods are also used to inhibit T cell
 CC activation in humans.

XX Sequence 8 AA;

Query Match 66.7%; Score 40; DB 21; Length 8;

Best Local Similarity 100.0%; Pred. No. 7.8e+05;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AEGDITOR 9
 | | | | | | | |
 DB 1 AEGDITOR 8

RESULT 5
 AAR98509
 ID AAR98509 standard; peptide: 10 AA.

XX AAR98509;

XX 04-MAR-1997 (first entry)

XX CD8 antagonist #4.

XX CD8 antagonist; inhibitor; T-cell activation; human; CD8 alpha chain;

KW thymic differentiation; transplantation; bone marrow; liver; heart; lung;

KW kidney; cornea; skin graft; graft versus host disease; therapy.

XX Synthetic.

XX WO9622106-A1.

XX 25-JUL-1996.

XX 17-JAN-1996; 96WO-US00310.

XX 17-JAN-1995; 95US-0372952.

XX (UYJE-) UNIV JEFFERSON THOMAS.

XX Choksi S, Huang Z, Jameson BA, Korgold R;

XX WPI; 1996-354307/35.

XX CD8 antagonist peptide(s) - used for inhibiting T cell activation,

PT partic. for treating transplant rejection or graft versus host

PT disease

XX Claim 16; Page 29; 44pp; English.

XX AAR98506-R98511 represent CD8 antagonists of the invention. CD8 plays a
 CC major role in the activation of mature T-cells, and in the thymic
 CC differentiation process that leads to CD8 expression. CD8 is expressed
 CC either as a homodimer (containing two alpha chains) or as a heterodimer
 CC (an alpha and a beta chain). The CD8 CD82-like region is involved in
 CC regulating T-cell activation. The antagonists of the invention comprise
 CC a molecular surface similar to at least a portion of human CD8 molecular
 CC surface around the site of one of these peptides. The compounds of the
 CC invention compete with CD8 so as to inhibit T-cell activation. They can
 CC also be used to treat an individual who is about to undergo, is
 CC undergoing, or has undergone a transplantation procedure such as bone
 CC marrow, liver, heart, kidney, lung, islets, or cornea transplantation,
 CC or skin grafts. The compounds can also be used to treat an individual
 CC suspected of suffering from, or susceptible to graft versus host
 CC disease.

XX Sequence 10 AA;

Query Match 66.7%; Score 40; DB 17; Length 10;

Best Local Similarity 100.0%; Pred. No. 0.15;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AEGDITOR 9
 | | | | | | | |
 DB 2 AEGDITOR 9

RESULT 6

AAY43649
 ID AAY43649 standard; peptide: 10 AA.

XX AAY43649;

XX 11-FEB-2000 (first entry)

XX Amino acid sequence of a human CD8 antagonist peptide.

KM Extracellular domain; alpha chain; human CD8; T cell activation;
 KM cytotoxic T lymphocyte activation; CTL activation;
 KM graft versus host disease; CD8 antagonist; transplant;
 KM allogeneic bone marrow transplant; cyclic.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FT Disulfide bond 1..10
 XX
 XX W0954345-A1.
 XX
 XX 28-OCT-1999.
 XX
 XX 21-APR-1999; 99WO-US08814.
 XX
 XX 21-APR-1998; 98US-0082436.
 XX
 XX (UYJE-) UNIV JEFFERSON THOMAS.
 XX
 XX Krongold R, Huang Z, Choksi S;
 DR WPI; 2000-013224/01.
 XX
 XX Novel CD8 antagonists used to inhibit cytotoxic T lymphocytes and
 PT prevent immune responses -
 XX
 XX Disclosure; Page 49; 67pp; English.
 XX
 XX The present sequence represents a human CD8 antagonist peptide of the
 CC invention. The specification describes peptides that are CD8 antagonists,
 CC and that compete with regions within the extracellular domain of the
 CC alpha chain of CD8 regions in intermolecular interactions that involve
 CC CD8 which are associated with cytotoxic T lymphocyte (CTL) activation.
 CC These peptides comprise a molecular surface that corresponds to a
 CC molecular surface of human CD8 at amino acids 73-76, 38-46, 53-56,
 CC 60-67 or 53-67. The peptides can be used in methods and compositions
 CC for therapeutically or prophylactically treating graft versus host
 CC disease in patients who have or about to undergo transplants, e.g.
 CC allogeneic bone marrow transplants. The methods and compositions are
 CC also used to treat individuals who are undergoing and/or who have
 CC undergone transplants, e.g. transplants of heart, liver, lung, islets
 CC or cornea, and skin grafts. The methods are also used to inhibit T cell
 CC activation in humans.
 CC
 XX
 XX Sequence 10 AA:
 SQ
 Query Match 66.7%; Score 40; DB 21; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.15;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AEGIDTOR 9
 | | | | | | |
 Db 2 AEGIDTOR 9

RESULT 7
 AAM97622
 ID AAM97622 standard; Peptide: 11 AA.
 XX
 AC AAM97622;
 XX
 DT 24-JAN-2002 (first entry)
 XX
 XX Human peptide #897 encoded by a SNP oligonucleotide.
 XX
 XX Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
 KM neuroprotective; antiinfectious; gene therapy; vaccine; amylose; cancer;
 KM amyloid protein; angiotensin; apoptosis related protein; cadherin;
 KM cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
 KM complement related protein; cytochrome; kinesin; cytokine; interferon;
 KM Interleukin; G-protein coupled receptor; thioesterase; inflammation;

KM multifactorial disease; autoimmune disease; infection;
 KM nervous system disease.
 XX
 OS Homo sapiens.
 XX
 XX W0200147944-A2.
 XX
 XX 05-JUL-2001.
 XX
 XX 28-DEC-2000; 2000WO-US35498.
 XX
 XX 28-DEC-1999; 99US-0173419.
 PR 27-DEC-2000; 2000US-0173419.
 XX
 XX (CURA-) CURAGEN CORP.
 XX
 XX Shinkels RA, Leach M;
 DR WPI; 2001-465210/50.
 XX
 XX Disclosure; Page 3864; 4143pp; English.
 XX
 XX The present invention relates to oligonucleotides (see AML26793-AML34659)
 CC encoding polymorphic variants of proteins related to amylases, amyloid
 CC proteins, angiotensin, apoptosis related proteins, cadherin, cyclin,
 CC polymerase, oncogenes, histones, kinases, colony stimulating factors,
 CC complement related proteins, cytochromes, kinesins, cytokines,
 CC interferons, interleukins, G-protein coupled receptors and thioesterases.
 CC The present sequence is a peptide encoded by one such oligonucleotide.
 CC The oligonucleotides and the peptides encoded by them may be used in the
 CC prevention, diagnosis and treatment of diseases associated with
 CC inappropriate expression of the proteins listed above. Disorders that may
 CC be prevented, diagnosed and/or treated include multifactorial diseases
 CC with a genetic component, such as autoimmune diseases (e.g. rheumatoid
 CC arthritis, multiple sclerosis, diabetes, systemic lupus erythematosus,
 CC and Grave's disease), inflammation, cancer (e.g. cancers of the bladder,
 CC brain, breast, colon and kidney, leukaemia), diseases of the nervous
 CC system and an infection of pathogenic organisms.
 CC
 XX
 XX Sequence 11 AA:
 SQ
 Query Match 53.3%; Score 32; DB 22; Length 11;
 Best Local Similarity 75.0%; Pred. No. 7.7;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 AEGIDTOR 9
 | | | | | | |
 Db 3 ARGVDTOR 10

RESULT 8
 ABG60421
 ID ABG60421 standard; Peptide: 9 AA.
 XX
 AC ABG60421;
 XX
 DT 30-JUL-2002 (first entry)
 XX
 XX Selective targeting peptide #96.
 XX
 XX Targeting peptide; cancer; arthritis; diabetes; inflammatory disease;
 KM atherosclerosis; autoimmune disease; bacterial infection; apoptosis;
 KM viral infection; cardiovascular disease; degenerative disease; ischaemia;
 KM inflammation; macular degeneration; antiinflammatory; antidiabetic;
 KM cardiovascular; immunomodulator; antibacterial; antiviral; cytostatic;
 KM gene therapy.
 XX
 OS Synthetic.

```

PN WO200220769-A1.
XX
PD 14-MAR-2002.
XX
PF 07-SEP-2001; 2001WO-US27692.
XX
PR 08-SEP-2000; 2000US-231266P.
XX 17-JAN-2001; 2001US-0765101.
XX (TEXA ) UNIV TEXAS SYSTEM.
XX
XX Arap W, Pasqualini R;
XX
XX MPI; 2002-415731/44.
XX
XX Targeting peptides identified by phage display, useful for targeting
XX delivery to an organ or tissue, particularly for treating a disease,
XX e.g. cancer, inflammatory or autoimmune diseases, infections or
XX cardiovascular disease -
XX
XX Claim 22: Page 89; 317pp; English.
XX
XX The invention relates to an isolated peptide of 100 amino acids or less
XX in size useful for targeting delivery to an organ or tissue, particularly
XX for treating a disease, e.g. cancer, arthritis, diabetes, inflammatory
XX disease, atherosclerosis, autoimmune disease, bacterial infection, viral
XX infection, cardiovascular disease or degenerative disease. The peptide is
XX also useful for inducing apoptosis, particularly to a subject with
XX ischemia, cancer, arthritis, diabetes, cardiovascular disease,
XX inflammation or macular degeneration. Furthermore, the peptide is useful
XX for diagnosing the diseases cited above. Targeting peptides of the
XX invention can also be used to deliver an agent to a foetus, by attaching
XX a peptide to the agent and administering the peptide to a pregnant
XX subject. Sequences ABG60326-ABG60574 represent selective targeting
XX peptides of the invention.
XX
XX Sequence 9 AA:
SQ
XX
XX Query Match 43.3%; Score 26; DB 23; Length 9;
XX Best Local Similarity 75.0%; Pred. No. 7.8e+05;
XX Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 4 GLDYGQRFSS 11
XX |||||
XX 1 GLDYGQRFSS 8
XX
XX RESULT 9
XX ABB78736
XX ID ABB78736 standard; Peptide: 11 AA.
XX
XX AC ABB78736;
XX
XX 22-JUN-2002 (first entry)
XX
XX Chicken calyntenin-1 internal peptide SEQ ID NO:13.
XX
XX Human: calyntenin-1; calyntenin-2; calyntenin-3; nervous system;
XX calcium binding protein; neuroprotective; antiinflammatory; nootropic;
XX anticonvulsant; cerebroprotective; cytosolic; ophthalmological; tumour;
XX analgesic; neuroleptic; vaccine; gene therapy; nervous system disorder;
XX metastasis; ARP2/3 complex; neovascularogenesis; neurodegenerative disease;
XX neuroinflammatory disease; epileptic seizure; retinal disease;
XX pathological pain syndrome; psychiatric disorder.
XX
XX Gallus gallus.
XX
XX WO200222819-A2.
XX
XX 21-MAR-2002.
XX
XX 13-SEP-2001; 2001WO-1B01662.
XX

```

```

PR 14-SEP-2000; 2000EP-0810830.
XX
XX (UYZU-) UNIV ZUERICH.
XX
XX Sonderegger P, Hintsch G, Kinter J, Meskenaitė V, Schrimpf S;
XX Vogt L, Zurlinden A;
XX
XX MPI; 2002-401904/43.
XX
XX Isolated nervous system calcium binding protein, selected from
XX calyntenin-1-3, useful as valuable agents for the treatment of
XX disorders of nervous system and in the development of drugs -
XX
XX Example 2; Page 52; 158pp; English.
XX
XX The present invention describes an isolated nervous system calcium
XX binding protein (I), selected from calyntenin-1, calyntenin-2 or
XX calyntenin-3, used as a pharmaceutical, having calcium binding activity
XX and/or capable of binding ARP2/3 complex. (I) has neuroprotective,
XX antiinflammatory, nootropic, anticonvulsant, cerebroprotective,
XX cytosolic, ophthalmological, analgesic and neuroleptic activities. (I)
XX and the polynucleotide encoding it (II) can be used in vaccines and in
XX gene therapy. (I) and (II) are useful for the screening and for the
XX preparation of a medicament for the treatment of disorders, in particular
XX disorders of nervous system, particularly central nervous system
XX including brain. (I) and (II) are also useful for the preparation of a
XX medicament for the treatment of tumours including prevention or reduction
XX of growth, expansion infiltration and metastasis of primary and
XX metastatic tumours, in particular brain tumour or tumours of retina,
XX where the tumours involve an enhanced activity of ARP2/3 complex or
XX protease functionally connected with (I), in their growth, expansion,
XX infiltration, metastasis and promotion of blood vessels or
XX neovascularogenesis. (I) and (II) are also useful for treating, preventing or
XX ameliorating negative effects of neurodegenerative diseases or
XX neuroinflammatory diseases or epileptic seizures, and for treating,
XX ameliorating or preventing retinal diseases, pathological pain syndromes,
XX psychiatric disorders, learning and memory functions in healthy persons,
XX and for treating tumours. The present sequence represents a chicken
XX calyntenin-1 peptide, which is used in an example from the present
XX invention.
XX
XX Sequence 11 AA:
SQ
XX
XX Query Match 43.3%; Score 26; DB 23; Length 11;
XX Best Local Similarity 60.0%; Pred. No. 1.3e+02;
XX Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
XX
XX QY 3 EGLDYGQRFSSG 12
XX |||||
XX 1 EGLDYGQRFSSG 10
XX
XX RESULT 10
XX AAM49400
XX ID AAM49400 standard; peptide: 9 AA.
XX
XX AC AAM49400;
XX
XX 05-JUN-1998 (first entry)
XX
XX Human leucocyte antigen DQ4 binding peptide #291.
XX
XX Human leucocyte antigen; HLA-DQ4; combinatorial library;
XX autoimmune disease; chronic articular rheumatism.
XX
XX Synthetic.
XX
XX JP08151396-A.
XX
XX 11-JUN-1996.
XX
XX 28-NOV-1994; 94JP-0292657.
XX

```

PR 28-NOV-1994; 94JP-0292657.
 XX
 PA (TEIJ) TEIJIN LTD.
 XX
 DR WPI; 1996-329479/33.
 XX
 PT HLA-binding oligopeptide and an immuno:regulator contg it - used in
 PT the treatment of auto:immune disease
 XX
 PS Claim 4; Page 36; 61pp; Japanese.
 XX
 CC This peptide is an example of a peptide which binds to a human leucocyte
 CC antigen HLA-P04 molecule. The peptide was isolated from a phagemid
 CC combinatorial library comprising the sequence AAV05953, by screening
 CC with an HLA-P04 molecule. The peptide is used for the treatment of
 CC autoimmune disease, or especially for treatment of viral diseases.
 XX
 SQ Sequence 9 AA;
 Query Match 41.7%; Score 25; DB 17; Length 9;
 Best Local Similarity 71.4%; Pred. No. 7.8e+05;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 OY 3 EGDTRQ 9
 DB 3 EGDTRQ 9
 RESULT 11
 AA015769
 ID AA015769 standard; Peptide; 11 AA.
 XX
 AC AA015769;
 XX
 DT 31-OCT-1997 (first entry)
 XX
 DE Anion exchange standard peptide.
 XX
 KW Coiled-coil heterodimer; fusion protein; purification; detection;
 KW coil peptide; peptide 0994; affinity chromatography;
 KW anion exchange.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1
 FT /note= "N-terminal Ac"
 FT Modified-site 11
 FT /note= "C-terminal amide"
 FT
 XX
 PN W09712988-A1.
 XX
 PD 10-APR-1997.
 XX
 PE 04-OCT-1996; 96WO-US16032.
 XX
 PR 06-OCT-1995; 95US-0540397.
 XX
 PA (PENEC-) PENEC.
 XX
 PI Bautista D, Cachia PJ, Hodges RS, Houston ME, Irwin RT;
 PI Tripet B, Yu L;
 XX
 DR WPI; 1997-226231/20.
 XX
 PT Nucleic acid segment encoding one subunit of a coiled dimer protein
 PT - also derived fusion proteins that can be detected or purified by
 PT reaction with the complementary second subunit
 XX
 PS Example 8; Page 36; 80pp; English.
 XX
 CC This peptide sequence comprises an anion exchange standard that
 CC has a net charge at pH 6.0 of -4. The peptide was used in

CC experiments to determine conditions for removal of non-specific
 CC peptides from a selective dimethylammonium affinity column carrying
 CC coil peptide 0994 (see AA015767). It was eluted from the column
 CC by washing with a solution of 1.0 M KCl/10 mM phosphate for 5 min.
 CC Peptide 0994 is used in methods for the detection and purification
 CC of expressed proteins via alpha-helical coiled-coil heterodimer
 CC formation.
 XX
 SQ Sequence 11 AA;
 Query Match 41.7%; Score 25; DB 18; Length 11;
 Best Local Similarity 83.3%; Pred. No. 2.2e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OY 1 AAEGLD 6
 DB 6 AAEGLD 11
 RESULT 12
 AA003264
 ID AA003264 standard; Peptide; 11 AA.
 XX
 AC AA003264;
 XX
 DT 12-SEP-2001 (first entry)
 XX
 DE Fruit fly G protein coupled receptors, DmGPCR6aL/DL ligand #13.
 XX
 KW Fruit fly; G protein coupled receptor; DmGPCR6aL/DL;
 KW human immunodeficiency virus; HIV; cancer; Parkinson's disease;
 KW diabetes; obesity; atherosclerosis; thrombosis; stroke; renal failure;
 KW inflammation; rheumatoid arthritis; autoimmune disorder;
 KW neurological disorder; schizophrenia; manic depression; dementia;
 KW severe mental retardation; dyskinesia; Huntington's disease;
 KW Tourette's syndrome; ligand.
 XX
 OS Drosophila melanogaster.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 11
 FT /note= "C-terminus is amidated"
 FT
 XX
 PN W0200131005-A2.
 XX
 PD 03-MAY-2001.
 XX
 PE 20-OCT-2000; 2000WO-US29002.
 XX
 PR 22-OCT-1999; 99US-0425676.
 XX
 PA (PHNA) PHARMACIA & UPJOHN CO.
 XX
 PI Lowery DE, Smith VG, Kudlak TA, Larsen MJ;
 PI WPI; 2001-316333/33.
 XX
 DR New Drosophila melanogaster GPCR nucleic acids and polypeptide useful
 DR for inducing an immune response, for identifying homologs and for
 DR treating e.g. diabetes, obesity and manic depression -
 XX
 PS Example 9; Page 99; 110pp; English.
 XX
 CC The sequence is a fruit fly G protein coupled receptors, DmGPCR6aL/DL,
 CC peptide ligand. The proteins are useful for inducing an immune response
 CC against itself in a mammal. The nucleic acids are useful for identifying
 CC an animal homolog of DmGPCR, by screening databases or libraries. The
 CC compounds identified as binding partners or modulators of GPCR binding
 CC are useful for treating diseases in animals, and for control insects that
 CC are harmful or cause injury to plants or animals. Diseases treated
 CC include infections (e.g. viral and human immunodeficiency virus, HIV),
 CC cancer, pain, Parkinson's disease, hypotension, hypertension, diabetes,
 CC obesity, atherosclerosis, thrombosis, stroke, renal failure,

CC inflammation, rheumatoid arthritis, autoimmune disorders, and psychotic
 CC and neurological disorders (anxiety, schizophrenia, manic depression,
 CC delirium, dementia, severe mental retardation, dyskinesias, Huntington's
 CC disease or Tourette's syndrome). The nucleic acids can be used for
 CC genetic mapping, and producing the GPCRs. Anti-GPCR antibodies can be
 CC used in therapy, diagnostic assays and for modulating GPCR activity.

XX
 SQ Sequence 11 AA:

Query Match 41.7%; Score 25; DB 22; Length 11;
 Best Local Similarity 60.0%; Pred. No. 2.2e+02;
 Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AEGIDTQRF 10
 | | | | |
 Db 2 AEEPLGTMRF 11

RESULT 13

AAU03828
 ID AAU03828 standard; Peptide: 11 AA.

XX
 AC AAU03828;

XX
 DT 12-SEP-2001 (first entry)

XX
 DE G protein-coupled receptor-like (GPCR-like) receptor binding peptide #5.

XX
 KW G protein-coupled receptor-like receptor; GPCR-like receptor; helminth;

XX
 KW endoparasite; ectoparasite; invertebrate; insect; neurological disorder;

XX
 KW neuromuscular disorder; human; nematode; lobster; locust; mollusc; leech;

XX
 KW fruitfly.

XX
 OS Caenorhabditis elegans.

XX
 PN W0200138533-A2.

XX
 PD 31-MAY-2001.

XX
 PF 24-NOV-2000; 2000WO-US32225.

XX
 PR 24-NOV-1999; 99US-0167523.

XX
 PA (PMAA) PHARMACIA & UPJOHN.

XX
 PI Lowery DE, Geary TG, Kubiak TM, Larsen MJ;

XX
 DR WPI; 2001-343952/36.

XX
 XX Using G-protein-coupled receptor (GPCR)-like receptors to identify

XX
 PT candidate compounds for the treatment and prevention of invertebrate

XX
 PT parasites, especially helminths and insects -

XX
 PS Claim 18; Page 58; 219pp; English.

XX
 CC The sequence represents a G protein-coupled receptor-like (GPCR-like)

XX
 CC receptor protein. GPCR-like receptors and their associated nucleic acids

XX
 CC may be used to identify candidate compounds for their ability to modulate

XX
 CC the activity of GPCRs. The sequences therefore are useful for treating

XX
 CC and preventing infection by endoparasitic and ectoparasitic invertebrate

XX
 CC parasites, especially helminths and insects, and particularly ailments

XX
 CC related to aberrant neurological and neuromuscular function.

XX
 SQ Sequence 11 AA:

Query Match 41.7%; Score 25; DB 22; Length 11;
 Best Local Similarity 60.0%; Pred. No. 2.2e+02;

Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AEGIDTQRF 10
 | | | | |
 Db 2 AEEPLGTMRF 11

RESULT 14

AA66812
 ID AA66812 standard; Peptide: 12 AA.

XX
 AC AA66812;

XX
 DT 11-SEP-1995 (first entry)

XX
 DE Syndecan heparan-attachment sequence used for fusion protein.

XX
 KW Cell surface; proteoglycan; syndecan; mouse; mammary; epithelial cell;

XX
 KW ectodomain; NMuMG; glycosylation; heparan sulphate; chondroitin sulphate;

XX
 KW chimaera; chimaeric molecule; effector molecule; receptor; drug;

XX
 KW antibody; diagnostic agent.

XX
 OS Mus musculus.

XX
 PN W09500633-A.

XX
 PD 05-JAN-1995.

XX
 PF 17-JUN-1994; 94MO-US06920.

XX
 PR 17-JUN-1993; 93US-0078683.

XX
 PA (CHIL-) CHILDRENS MEDICAL CENT.

XX
 PA (STRD) UNIV LELAND STANFORD JUNIOR.

XX
 PI Bernfield M, Kato M, Saunders S;

XX
 DR WPI; 1995-052071/07.

XX
 PT DNA and protein sequences for recombinant syndecan-derived

XX
 PT proteoglycans - comprising a core protein having glycosylation

XX
 PT sites for heparan sulphate glycosaminoglycan side chains.

XX
 PS Claim 25; Page 84; 97pp; English.

XX
 CC An example of a heparan sulphate attachment site derived from syndecan

XX
 CC proteins, used in the construction of a fusion protein. The sequence is

XX
 CC based on the generic attachment site (AA66811) derived from amino acid

XX
 CC sequence comparisons between mouse, rat, hamster and human homologs of

XX
 CC the syndecan-1 protein sequence. The glycosylation site is found in the

XX
 CC extracellular domain of the protein. The fusion protein contains the

XX
 CC heparan attachment site, a further chondroitin sulphate

XX
 CC glycosaminoglycan group linked to biological effector molecules, cell

XX
 CC surface receptors, drugs, antibodies, diagnostic agents or components of

XX
 CC microorganisms.

XX
 SQ Sequence 12 AA:

Query Match 41.7%; Score 25; DB 16; Length 12;
 Best Local Similarity 55.6%; Pred. No. 2.4e+02;

Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 GLDTRFSG 12
 | | | | |
 Db 2 GDSNDNMSG 10

RESULT 15

AA61860
 ID AA61860 standard; Peptide: 9 AA.

XX
 AC AA61860;

XX
 DT 18-MAY-1995 (first entry)

XX
 DE MBP peptide 108, potential binder of HLA-A2.1.

XX
 KW antigen; epitope; immunogenic target protein; PSA; HBVC; HBVS; EBV;

XX
 KW HIV1; plasma specific antigen; hepatitis B virus; Epstein Barr;

KW human immunodeficiency virus; human papilloma virus; p53; c-erbB2;
 KW MAGE-1; melanoma antigen-1; core antigen; surface antigen;
 KW pharmaceutical composition; in vivo; ex vivo; therapeutic;
 KW diagnostic; MHC class I molecule; major histocompatibility complex;
 KW HLA-A2.1; 9mer; 10mer; anchor; human leukocyte antigen; PLP; 6mer;
 KW algorithm prediction; MBP.
 XX
 OS Homo sapiens.
 XX
 PN WO9420127-A.
 XX
 PD 15-SEP-1994.
 XX
 PE 04-MAR-1994; 94WO-US02353.
 XX
 PR 05-MAR-1993; 93US-0027146.
 PR 04-JUN-1993; 93US-0073205.
 PR 29-NOV-1993; 93US-0159184.
 XX
 PA (CYTE-) CYTEL CORP.
 XX
 PI Grey HM, Kast WM, Sette A, Sidney J;
 XX
 DR WPI: 1994-302678/37.
 XX
 PT Immunogenic peptide(s) having an HLA-A2.1 binding motif - used
 PT for treatment or prophylaxis of cancer, virus infection or
 PT autoimmune diseases.
 XX
 PS Disclosure: Page 122; 138pp; English.
 XX
 CC AAR61714-837 are potential peptide binders of HLA-A2.1 motif. These
 CC peptides are thus potentially immunogenic. They were predicted by
 CC using an algorithm, which assigns a score for each amino acid, at
 CC each position along a peptide. A peptide is scored in the 'Grouped
 CC Ratio' algorithm as a product of the scores of each of its residues.
 CC This peptide has an algorithm score (EO2) of -23.34. This value can
 CC then be used to predict a population of peptides with the highest
 CC occurrence of good binders. The peptides of the invention can induce
 CC cytotoxic T lymphocytes which can react with target cells. They can
 CC be used for the treatment or prophylaxis of cancer, eg. prostate
 CC cancer or lymphoma, etc.
 CC
 SO Sequence 9 AA;
 Query Match 40.0%; Score 24; DB 15; Length 9;
 Best Local Similarity 62.5%; Pred. No. 7.8e+05;
 Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 4 GLDTRFS 11
 Db 1 GLSISRFS 8

Search completed: April 15, 2003, 10:20:06
 Job time : 36 secs

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 15, 2003, 10:21:05 ; Search time 14 Seconds

(without alignments)
52.402 Million cell updates/sec

Title: US-09-521-527C-1

Perfect score: 60

Sequence: 1 AAEGCDTQRRSG 12

Scoring table: BLOSUM62

Searched: 248812 seqs, 61136040 residues

Total number of hits satisfying chosen parameters: 34263

Minimum DB seq length: 0

Maximum DB seq length: 12

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Published_Applications_AA:*

1: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubppaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep:*
5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubppaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep:*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22	36.7	8	9	US-10-079-625-33
2	22	36.7	9	10	US-09-753-831-10
3	22	36.7	9	10	US-09-753-831-28
4	22	36.7	10	10	US-09-920-174-35
5	21	35.0	5	10	US-09-859-214-18
6	21	35.0	5	10	US-09-859-214-19
7	21	35.0	6	8	US-08-875-849C-14
8	21	35.0	6	9	US-10-213-539-1
9	21	35.0	6	10	US-09-859-214-13
10	21	35.0	6	10	US-09-859-214-15
11	21	35.0	6	10	US-09-859-214-17
12	21	35.0	7	10	US-09-859-214-9
13	21	35.0	7	10	US-09-859-214-12
14	21	35.0	7	10	US-09-859-214-14
15	21	35.0	7	10	US-09-859-214-16
16	21	35.0	8	10	US-09-859-214-58
17	21	35.0	9	10	US-09-859-214-11
18	21	35.0	9	10	US-09-834-765-82
19	21	35.0	10	10	US-09-824-206-5

20	21	35.0	10	10	US-09-859-214-55	Sequence 55, Appl
21	21	35.0	10	10	US-09-859-214-59	Sequence 59, Appl
22	21	35.0	10	10	US-09-147-490-1	Sequence 1, Appl
23	21	35.0	10	10	US-09-779-308-37	Sequence 37, Appl
24	21	35.0	11	10	US-09-950-313-38	Sequence 30, Appl
25	21	35.0	12	9	US-09-826-290-82	Sequence 82, Appl
26	21	35.0	12	10	US-09-859-214-56	Sequence 56, Appl
27	21	35.0	12	10	US-09-859-214-60	Sequence 60, Appl
28	21	35.0	12	12	US-10-090-624-22	Sequence 22, Appl
29	21	35.0	12	12	US-10-052-798-14	Sequence 14, Appl
30	20	33.3	7	9	US-09-281-485-25	Sequence 25, Appl
31	20	33.3	8	10	US-09-984-056-54	Sequence 54, Appl
32	20	33.3	8	10	US-09-984-057-54	Sequence 54, Appl
33	20	33.3	9	9	US-09-931-325A-156	Sequence 156, App
34	19	31.7	6	10	US-09-867-852-133	Sequence 133, App
35	19	31.7	7	10	US-09-832-723-11	Sequence 11, Appl
36	19	31.7	8	9	US-10-014-485A-53	Sequence 53, Appl
37	19	31.7	8	9	US-09-910-033A-6	Sequence 6, Appl
38	19	31.7	9	9	US-10-102-283-119	Sequence 119, App
39	19	31.7	9	9	US-10-102-283-138	Sequence 138, App
40	19	31.7	9	9	US-10-091-135-11	Sequence 11, Appl
41	19	31.7	9	9	US-10-125-635A-85	Sequence 85, Appl
42	19	31.7	9	9	US-10-125-635A-235	Sequence 235, App
43	19	31.7	9	9	US-10-125-635A-302	Sequence 302, App
44	19	31.7	9	9	US-09-931-325A-157	Sequence 157, App
45	19	31.7	9	10	US-09-834-765-450	Sequence 450, App

ALIGNMENTS

RESULT 1
US-10-079-625-33
Sequence 33, Application US/10079625
Publication No. US20020182676A1
GENERAL INFORMATION:
APPLICANT: Tartaglia, Louis A.
APPLICANT: Tepper, Robert I.
APPLICANT: Culpepper, Janice A.
APPLICANT: White, David W.
TITLE OF INVENTION: THE OB RECEPTOR AND METHODS FOR
TITLE OF INVENTION: THE DIAGNOSIS AND TREATMENT OF BODY WEIGHT DISORDERS,
TITLE OF INVENTION: INCLUDING OBESITY AND CACHEXIA
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson, P.C.
STREET: 225 Franklin Street
City: Boston
STATE: MA
COUNTRY: US
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/079,625
FILING DATE: 2002-FEB-19
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/864,564
FILING DATE: 28-MAY-1997
APPLICATION NUMBER: 08/708,123
FILING DATE: 03-SEP-1996
APPLICATION NUMBER: 08/638,524
FILING DATE: 26-APR-1996
APPLICATION NUMBER: 08/599,455
FILING DATE: 22-JAN-1996
APPLICATION NUMBER: 08/583,153
FILING DATE: 28-DEC-1995
APPLICATION NUMBER: 08/570,142
FILING DATE: 11-DEC-1995
APPLICATION NUMBER: 08/569,485

FILE DATE: 08-DEC-1995
APPLICATION NUMBER: 08/566,622
FILING DATE: 04-DEC-1995
APPLICATION NUMBER: 08/562,663
FILING DATE: 27-NOV-1995
ATTORNEY/AGENT INFORMATION:
NAME: Melkiojohn, Ph.D., Anita L.
REGISTRATION NUMBER: 35,283
REFERENCE/DOCKET NUMBER: 07334/019002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-5070
TELEFAX: 617-542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-10-079-625-33

Query Match 36.7%; Score 22; DB 9; Length 8;
Best Local Similarity 50.0%; Pred. No. 2.2e+05;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 AEGIDTQ 9
|:|:|:|:
DB 1 AQCINPQK 8

RESULT 2
US-09-753-831-10
Sequence 10, Application US/09753831
Patent No. US20020137683A1
GENERAL INFORMATION:
APPLICANT: Hogan, Kevin T.
TITLE OF INVENTION: C-SKI Oncogene-Derived Peptides for Prevention,
FILE REFERENCE: 26747-27
CURRENT APPLICATION NUMBER: US/09/753,831
CURRENT FILING DATE: 2001-01-03
PRIOR APPLICATION NUMBER: U.S. 60/174296
PRIOR FILING DATE: 2000-01-03
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 10
LENGTH: 9
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: No. US20020137683A1apeptide
OTHER INFORMATION: derived from the human C-SKI oncoprotein.
US-09-753-831-10

Query Match 36.7%; Score 22; DB 10; Length 9;
Best Local Similarity 80.0%; Pred. No. 2.2e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 GLDTQ 8
|:|:|:|:
DB 5 GLDTK 9

RESULT 3
US-09-753-831-28
Sequence 28, Application US/09753831
Patent No. US20020137683A1
GENERAL INFORMATION:
APPLICANT: Hogan, Kevin T.
APPLICANT: Ross, Mark M.
TITLE OF INVENTION: C-SKI Oncogene-Derived Peptides for Prevention,
TITLE OF INVENTION: Treatment and Diagnosis of Cancer

FILE REFERENCE: 26747-27
CURRENT APPLICATION NUMBER: US/09/753,831
CURRENT FILING DATE: 2001-01-03
PRIOR APPLICATION NUMBER: U.S. 60/174296
PRIOR FILING DATE: 2000-01-03
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 28
LENGTH: 9
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: No. US20020137683A1apeptide
OTHER INFORMATION: derived from the human C-SKI oncoprotein.
US-09-753-831-28

Query Match 36.7%; Score 22; DB 10; Length 9;
Best Local Similarity 80.0%; Pred. No. 2.2e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 GLDTQ 8
|:|:|:|:
DB 5 GLDTK 9

RESULT 4
US-09-920-174-35
Sequence 35, Application US/09920174
Patent No. US20020150590A1
GENERAL INFORMATION:
APPLICANT: KHANNA, RAJIV
APPLICANT: KERR, BEVERLEY M.
APPLICANT: MISKO, THOR S.
APPLICANT: MOSS, DENIS J.
APPLICANT: BURROWS, SCOTT R.
TITLE OF INVENTION: EBV CTL EPITOPES
FILE REFERENCE: FBRC:008USC2
CURRENT APPLICATION NUMBER: US/09/920,174
CURRENT FILING DATE: 2001-08-01
PRIOR APPLICATION NUMBER: 09/920,175
PRIOR FILING DATE: 2001-08-01
PRIOR APPLICATION NUMBER: 09/194,450
PRIOR FILING DATE: 1998-11-23
NUMBER OF SEQ ID NOS: 50
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 35
LENGTH: 10
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
OTHER INFORMATION: Peptide
US-09-920-174-35

Query Match 36.7%; Score 22; DB 10; Length 10;
Best Local Similarity 62.5%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 LDTQRFSG 12
|:|:|:|:
DB 2 LDFVRFMG 9

RESULT 5
US-09-859-214-18
Sequence 18, Application US/09859214
Patent No. US20020103111A1
GENERAL INFORMATION:
APPLICANT: Schwender, Charles F.
APPLICANT: Shroff, Hitesh N.
TITLE OF INVENTION: INHIBITORS OF MAGCAM-1-MEDIATED
INTERACTIONS AND METHODS OF USE THEREFOR
NUMBER OF SEQUENCES: 89

;;
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
;; STREET: Two Millitia Drive
;; CITY: Lexington
;; STATE: Massachusetts
;; COUNTRY: USA
;; ZIP: 02421
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.30
;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/859,214
;; FILING DATE: 16-May-2001
;; CLASSIFICATION: <Unknown>
;;
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 09/109,879
;; FILING DATE: <Unknown>
;; APPLICATION NUMBER: US 08/582,740
;; FILING DATE: 04-JAN-1996
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Brook, David E.
;; REGISTRATION NUMBER: 22,592
;; REFERENCE/DOCKET NUMBER: LKS95-12A2
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (781) 861-6240
;; TELEFAX: (781) 861-9540
;;
;; INFORMATION FOR SEQ ID NO: 18:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 5 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: <Unknown>
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; FEATURE:
;; NAME/KEY: Modified-site
;; LOCATION: 1
;; OTHER INFORMATION: /label= modified aa
;; /note= "Ac - Glycine"
;;
;; NAME/KEY: Modified-site
;; LOCATION: 5
;; OTHER INFORMATION: /label= modified aa
;; /note= "Serine - NH2"
;; SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-09-859-214-18
Query Match 35.0%; Score 21; DB 10; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 GLDT 7
Db 1 GLDT 4
RESULT 6
US-09-859-214-19
; Sequence 19, Application US/09859214
; Patent No. US2002010311A1
; GENERAL INFORMATION:
; APPLICANT: Schwender, Charles F.
; Shroff, Hitesh N.
; TITLE OF INVENTION: INHIBITORS OF MADCAM-1-MEDIATED
; INTERACTIONS AND METHODS OF USE THEREFOR
; NUMBER OF SEQUENCES: 89
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Millitia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA

;;
;; ZIP: 02421
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.30
;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/859,214
;; FILING DATE: 16-May-2001
;; CLASSIFICATION: <Unknown>
;;
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 09/109,879
;; FILING DATE: <Unknown>
;; APPLICATION NUMBER: US 08/582,740
;; FILING DATE: 04-JAN-1996
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Brook, David E.
;; REGISTRATION NUMBER: 22,592
;; REFERENCE/DOCKET NUMBER: LKS95-12A2
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (781) 861-6240
;; TELEFAX: (781) 861-9540
;;
;; INFORMATION FOR SEQ ID NO: 19:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 5 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: <Unknown>
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; FEATURE:
;; NAME/KEY: Modified-site
;; LOCATION: 1
;; OTHER INFORMATION: /label= modified aa
;; /note= "Ac - Arginine"
;;
;; NAME/KEY: Modified-site
;; LOCATION: 5
;; OTHER INFORMATION: /label= modified aa
;; /note= "Threonine - NH2"
;; SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-09-859-214-19
Query Match 35.0%; Score 21; DB 10; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 GLDT 7
Db 2 GLDT 5
RESULT 7
US-08-875-849C-14
; Sequence 14, Application US/08875849C
; Patent No. US2002014731A1
; GENERAL INFORMATION:
; APPLICANT: Briskin, Michael J.
; Ringler, Douglas J.
; Picarelle, Dominic
; Newman, Walter
; TITLE OF INVENTION: Mucosal Vascular Addressins and Uses
; Thereof
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: 530 Virginia Road, PO Box 9133
; CITY: Concord
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 01742-9133
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/875,849C
FILING DATE: 08-Sep-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/02153
FILING DATE: 12-FEB-1996
APPLICATION NUMBER: US 08/523,004
FILING DATE: 01-SEP-1995
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: LKS94-04A2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 978-341-0036
TELEFAX: 978-341-0136
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-08-875-849c-14

Query Match 35.0%; Score 21; DB 8; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 GLDT 7
|||
Db 1 GLDT 4

RESULT 8
US-10-213-539-1
Sequence 1, Application US/10213539
Publication No. US20030003108A1
GENERAL INFORMATION:
APPLICANT: Fong, Sherman
APPLICANT: Hillan, Kenneth J.
TITLE OF INVENTION: Diagnosis and Treatment of Hepatic Disorders
FILE REFERENCE: P1081R1
CURRENT APPLICATION NUMBER: US/10/213,539
CURRENT FILING DATE: 2002-08-06
PRIOR APPLICATION NUMBER: US/09/310,247
PRIOR FILING DATE: 1999-05-12
PRIOR APPLICATION NUMBER: US 60/085,326
PRIOR FILING DATE: 1998-05-13
NUMBER OF SEQ ID NOS: 2
SEQ ID NO 1
LENGTH: 6
TYPE: PRT
ORGANISM: Homo sapiens
US-10-213-539-1

Query Match 35.0%; Score 21; DB 9; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 GLDT 7
|||
Db 1 GLDT 4

RESULT 9
US-09-859-214-13
Sequence 13, Application US/09859214
Patent No. US2002010311A1
GENERAL INFORMATION:
APPLICANT: Schwender, Charles F.

Shroff, Hitesh N.
TITLE OF INVENTION: INHIBITORS OF MACAM-1-MEDIATED
INTERACTIONS AND METHODS OF USE THEREFOR
NUMBER OF SEQUENCES: 89
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: USA
ZIP: 02421
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/859,214
FILING DATE: 16-May-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/109,879
FILING DATE: <Unknown>
APPLICATION NUMBER: US 08/582,740
FILING DATE: 04-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: LKS95-12A2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781) 861-6240
TELEFAX: (781) 861-9540
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /label= Modified aa
/note= "Ac - Tryptophan"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 6
OTHER INFORMATION: /label= Modified aa
/note= "Ileuonine - NH2"
SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-09-859-214-13

Query Match 35.0%; Score 21; DB 10; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 GLDT 7
|||
Db 3 GLDT 6

RESULT 10
US-09-859-214-15
Sequence 15, Application US/09859214
Patent No. US2002010311A1
GENERAL INFORMATION:
APPLICANT: Schwender, Charles F.
Shroff, Hitesh N.
TITLE OF INVENTION: INHIBITORS OF MACAM-1-MEDIATED
INTERACTIONS AND METHODS OF USE THEREFOR
NUMBER OF SEQUENCES: 89
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.

STREET: Two Millitia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: USA
ZIP: 02421

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/859,214
FILING DATE: 16-May-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/109,879
FILING DATE: 04-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: LKS95-12A2
TELEPHONE: (781) 861-9540
TELEFAX: (781) 861-6240

INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /label= Modified aa
/note= "Ac-Glycine"

FEATURE:
NAME/KEY: Modified-site
LOCATION: 6
OTHER INFORMATION: /label= Modified aa
/note= "Leucine - NH2"

SEQUENCE DESCRIPTION: SEQ ID NO: 15:
US-09-859-214-15

Query Match
Best Local Similarity 35.0%; Score 21; DB 10; Length 6;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GLDT 7
Db 1 GLDT 4

RESULT 11
US-09-859-214-17
Sequence 17, Application US/09859214
Patent No. US2002010311A1

GENERAL INFORMATION:
APPLICANT: Schwender, Charles F.
Shroff, Hitesh N.
TITLE OF INVENTION: INHIBITORS OF MADCAM-1-MEDIATED
INTERACTIONS AND METHODS OF USE THEREFOR
NUMBER OF SEQUENCES: 89
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millitia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: USA
ZIP: 02421

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/859,214
FILING DATE: 16-May-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/109,879
FILING DATE: 04-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: LKS95-12A2
TELEPHONE: (781) 861-9540
TELEFAX: (781) 861-6240

INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /label= modified aa
/note= "Ac - Arginine"

FEATURE:
NAME/KEY: Modified-site
LOCATION: 6
OTHER INFORMATION: /label= modified aa
/note= "Serine - NH2"

SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-09-859-214-17

Query Match
Best Local Similarity 35.0%; Score 21; DB 10; Length 6;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GLDT 7
Db 2 GLDT 5

RESULT 12
US-09-859-214-9
Sequence 9, Application US/09859214
Patent No. US2002010311A1

GENERAL INFORMATION:
APPLICANT: Schwender, Charles F.
Shroff, Hitesh N.
TITLE OF INVENTION: INHIBITORS OF MADCAM-1-MEDIATED
INTERACTIONS AND METHODS OF USE THEREFOR
NUMBER OF SEQUENCES: 89
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millitia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: USA
ZIP: 02421

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/859,214

FILED DATE: 16-May-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/109,879
FILING DATE: <Unknown>
APPLICATION NUMBER: US 08/582,740
FILING DATE: 04-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: LKS95-12A2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781) 861-6240
TELEFAX: (781) 861-9540
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /label= modified aa
/note= "Ac - Histidine"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 7
OTHER INFORMATION: /label= Modified aa
/note= "Threonine-NH2"
SEQUENCE DESCRIPTION: SEQ ID NO: 9:
US-09-859-214-9

Query Match 35.0%; Score 21; DB 10; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 GLDT 7
||||
Db 4 GLDT 7

RESULT 13
US-09-859-214-12
Sequence 12, Application US/09859214
Patent No. US2002010311A1
GENERAL INFORMATION:
APPLICANT: Schwender, Charles F.
Shroff, Hitesh N.
TITLE OF INVENTION: INHIBITORS OF MAGCAM-1-MEDIATED
INTERACTIONS AND METHODS OF USE THEREFOR
NUMBER OF SEQUENCES: 89
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millitia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: USA
ZIP: 02421
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/859,214
FILING DATE: 16-May-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/109,879
FILING DATE: <Unknown>
APPLICATION NUMBER: US 08/582,740

FILED DATE: 04-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: LKS95-12A2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781) 861-6240
TELEFAX: (781) 861-9540
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /label= modified aa
/note= "Ac - tryptophan"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 7
OTHER INFORMATION: /label= Modified aa
/note= "Serine - NH2"
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-859-214-12

Query Match 35.0%; Score 21; DB 10; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 GLDT 7
||||
Db 3 GLDT 6

RESULT 14
US-09-859-214-14
Sequence 14, Application US/09859214
Patent No. US2002010311A1
GENERAL INFORMATION:
APPLICANT: Schwender, Charles F.
Shroff, Hitesh N.
TITLE OF INVENTION: INHIBITORS OF MAGCAM-1-MEDIATED
INTERACTIONS AND METHODS OF USE THEREFOR
NUMBER OF SEQUENCES: 89
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millitia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: USA
ZIP: 02421
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/859,214
FILING DATE: 16-May-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/109,879
FILING DATE: <Unknown>
APPLICATION NUMBER: US 08/582,740
FILING DATE: 04-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: LKS95-12A2
TELECOMMUNICATION INFORMATION:

TELEPHONE: (781) 861-6240
TELEFAX: (781) 861-9540
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS: <unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /label= modified aa
/note= "Ac - Arginine"
NAME/KEY: Modified-site
LOCATION: 7
OTHER INFORMATION: /label= modified aa
/note= "Leucine - NH2"
SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-859-214-14

Query Match 35.0%; Score 21; DB 10; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GLDT 7
||||
DB 2 GLDT 5

RESULT 15
US-09-859-214-16
Sequence 16, Application US/09859214
Patent No. US2002010311A1
GENERAL INFORMATION:
APPLICANT: Schwender, Charles F.
Shoof, Hitesh N.
TITLE OF INVENTION: INHIBITORS OF MAGCAM-1-MEDIATED
INTERACTIONS AND METHODS OF USE THEREFOR
NUMBER OF SEQUENCES: 89
CORRESPONDENCE ADDRESS:
ADDRESSER: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: USA
ZIP: 02421
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/859,214
FILING DATE: 16-May-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/109,879
FILING DATE: <Unknown>
APPLICATION NUMBER: US 08/582,740
FILING DATE: 04-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: BROOK, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: LKS95-12A2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781) 861-9540
TELEFAX: (781) 861-6240
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid

STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /label= modified aa
/note= "Ac - Tryptophan"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 7
OTHER INFORMATION: /label= modified aa
/note= "Serine - NH2"
SEQUENCE DESCRIPTION: SEQ ID NO: 16:
US-09-859-214-16

Query Match 35.0%; Score 21; DB 10; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GLDT 7
||||
DB 3 GLDT 6

Search completed: April 15, 2003, 10:28:03
Job time : 15 secs

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 15, 2003, 10:19:29 ; Search time 14 Seconds
(without alignments)
25.220 Million cell updates/sec

Title: US-09-521-527C-1

Perfect score: 60

Sequence: 1 AAEGDTRRFG 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 91334

Minimum DB seq length: 0
Maximum DB seq length: 12

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents.AA:*
1: /cgn2_6/ptodata1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata1/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	66.7	8	1	US-08-372-952-3
2	40	66.7	8	1	US-08-875-309-3
3	40	66.7	8	5	PCT-US96-00310-3
4	40	66.7	10	1	US-08-372-952-6
5	40	66.7	10	4	US-08-875-309-6
6	40	66.7	10	5	PCT-US96-00310-6
7	26	43.3	12	4	US-09-242-435-16
8	23	38.3	8	4	US-09-187-859-1699
9	23	38.3	8	6	5168051-15
10	23	38.3	9	4	US-09-117-860-13
11	23	38.3	9	4	US-09-644-600-82
12	22	36.7	8	2	US-08-599-455B-33
13	22	36.7	8	4	US-09-069-781B-33
14	22	36.7	8	4	US-09-137-132-33
15	22	36.7	8	4	US-08-864-564A-33
16	22	36.7	8	4	US-09-094-410-33
17	22	36.7	12	1	US-08-036-555B-164
18	22	36.7	12	1	US-08-469-569-164
19	22	36.7	12	1	US-08-249-322A-164
20	22	36.7	12	1	US-08-469-526A-164
21	22	36.7	12	2	US-08-734-591A-164
22	22	36.7	12	2	US-08-469-660-164
23	22	36.7	12	2	US-08-466-860-5
24	22	36.7	12	4	US-08-472-040A-5
25	22	36.7	12	4	US-08-470-335-164
26	22	36.7	12	4	US-08-735-021-164
27	22	36.7	12	4	US-08-734-664A-164

28	22	36.7	12	4	US-08-276-776-5	Sequence 5, Appl1
29	22	36.7	12	4	US-08-471-209-5	Sequence 5, Appl1
30	22	36.7	12	4	US-08-470-339-164	Sequence 164, App
31	22	36.7	12	4	US-08-467-602-164	Sequence 164, App
32	22	36.7	12	5	PCT-US94-05083C-160	Sequence 160, App
33	22	36.7	12	5	PCT-US95-06846A-164	Sequence 164, App
34	21	35.0	5	3	US-08-582-740-18	Sequence 18, Appl
35	21	35.0	5	4	US-08-582-740-19	Sequence 19, Appl
36	21	35.0	5	4	US-09-109-879-18	Sequence 18, Appl
37	21	35.0	5	4	US-09-109-879-19	Sequence 19, Appl
38	21	35.0	6	3	US-08-582-740-13	Sequence 13, Appl
39	21	35.0	6	3	US-08-582-740-15	Sequence 15, Appl
40	21	35.0	6	3	US-08-582-740-17	Sequence 17, Appl
41	21	35.0	6	4	US-09-109-879-13	Sequence 13, Appl
42	21	35.0	6	4	US-09-109-879-15	Sequence 15, Appl
43	21	35.0	6	4	US-09-109-879-17	Sequence 17, Appl
44	21	35.0	7	2	US-08-627-173-11	Sequence 11, Appl
45	21	35.0	7	2	US-08-535-882A-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1
US-08-372-952-3

Sequence 3, Application US/08372952

Patent No. 5645837

GENERAL INFORMATION:

APPLICANT: Jameson, Bradford A.

APPLICANT: Choksi, Swati

APPLICANT: Koringold, Robert

TITLE OF INVENTION: CD8 Antagonists

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &

ADDRESSEE: No. 5645837ris

STREET: One Liberty Place, 46th Floor

CITY: Philadelphia

STATE: PA

COUNTRY: USA

ZIP: 19103

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/372,952

FILING DATE:

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Deluca, Mark

REGISTRATION NUMBER: 33,229

REFERENCE/DOCKET NUMBER: TJU-1440

TELECOMMUNICATION INFORMATION:

TELEPHONE: 215-568-3100

TELEFAX: 215-568-3100

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 8 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-372-952-3

Query Match 66.7%; Score 40; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.9e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 ASGLDTR 9
Db 1 ASGLDTR 8

RESULT 2

US-08-875-309-3
Sequence 3, Application US/08875309
Patent No. 6180600

GENERAL INFORMATION:

APPLICANT: Jameson, Bradford A.
APPLICANT: Choksi, Swati
APPLICANT: Korgold, Robert
APPLICANT: Huang, Zilwei
TITLE OF INVENTION: CD8 Antagonists
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & NO. 6180600ris
STREET: One Liberty Place, 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/875.309
FILING DATE: 26-NOV-1997

CLASSIFICATION: 424

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/00310
FILING DATE: 17-JAN-1996
CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/372.952
FILING DATE: 17-JAN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:

NAME: Deluca, Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: TJU-1772
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-875-309-3

Query Match 66.7%; Score 40; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.9e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 ASGLDTR 9
Db 1 ASGLDTR 8

RESULT 3

PCT-US96-00310-3
Sequence 3, Application PC/TUS9600310

GENERAL INFORMATION:

APPLICANT: Jameson, Bradford A.
APPLICANT: Choksi, Swati
APPLICANT: Korgold, Robert
APPLICANT: Huang, Zilwei
TITLE OF INVENTION: CD8 Antagonists

NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & Norris
STREET: One Liberty Place, 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/00310
FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/372.952
FILING DATE: 17-JAN-1995
CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Deluca, Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: TJU-1752
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:
LENGTH: 8 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US96-00310-3

Query Match 66.7%; Score 40; DB 5; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.9e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 ASGLDTR 9
Db 1 ASGLDTR 8

RESULT 4

US-08-372-952-6
Sequence 6, Application US/08372952

GENERAL INFORMATION:

APPLICANT: Jameson, Bradford A.
APPLICANT: Choksi, Swati
APPLICANT: Korgold, Robert
TITLE OF INVENTION: CD8 Antagonists
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
ADDRESSEE: No. 5645837ris
STREET: One Liberty Place, 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/372.952
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:

```

; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-1440
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-372-952-6

Query Match      66.7%; Score 40; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AEGDITOR 9
   |||||
Db 2 AEGDITOR 9

RESULT 5
US-08-875-309-6
; Sequence 6, Application US/08875309
; Patent No. 6180600
; GENERAL INFORMATION:
; APPLICANT: Jameson, Bradford A.
; APPLICANT: Choksi, Swati
; APPLICANT: Korngold, Robert
; APPLICANT: Huang, Zhiwei
; TITLE OF INVENTION: CD8 Antagonists
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 6180600rls
; STREET: One Liberty Place, 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/875,309
; FILING DATE: 26-NOV-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/00310
; FILING DATE: 17-JAN-1996
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/372,952
; FILING DATE: 17-JAN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-1772
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
```

```

; TOPOLOGY: circular
; MOLECULE TYPE: protein
; US-08-875-309-6

Query Match      66.7%; Score 40; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AEGDITOR 9
   |||||
Db 2 AEGDITOR 9

RESULT 6
PCT-US96-00310-6
; Sequence 6, Application PC/TUS9600310
; GENERAL INFORMATION:
; APPLICANT: Jameson, Bradford A.
; APPLICANT: Choksi, Swati
; APPLICANT: Korngold, Robert
; APPLICANT: Huang, Zhiwei
; TITLE OF INVENTION: CD8 Antagonists
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & Norris
; STREET: One Liberty Place, 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/00310
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/372,952
; FILING DATE: 17-JAN-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-1752
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: circular
; MOLECULE TYPE: protein
; PCT-US96-00310-6

Query Match      66.7%; Score 40; DB 5; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AEGDITOR 9
   |||||
Db 2 AEGDITOR 9

RESULT 7
US-09-242-435-16
; Sequence 16, Application US/09242435
; Patent No. 6461863
; GENERAL INFORMATION:
; APPLICANT: JARVIS, DONALD L.
```

```
; TITLE OF INVENTION: MODIFYING INSECT CELL GLYCOSYLATION PATHWAYS WITH
; FILE OF INVENTION: BACULOVIRUS EXPRESSION VECTORS
; FILE REFERENCE: UWYO:002US
; CURRENT APPLICATION NUMBER: US/09/242,435
; CURRENT FILING DATE: 1999-02-16
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 16
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; US-09-242-435-16

Query Match          43.3%; Score 26; DB 4; Length 12;
Best Local Similarity 71.4%; Pred. No. 59;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 DTORFSG 12
   1 :|||||
   1 DTSRFSG 7

Db

RESULT 8
; US-09-187-859-1699
; Sequence 1699, Application US/09187859A
; Patent No. 6358920
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; FILE REFERENCE: 100086.407C1
; CURRENT APPLICATION NUMBER: US/09/187,859A
; CURRENT FILING DATE: 1998-11-06
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1699
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Representative cyclic modulating agent based on
; OTHER INFORMATION: cadherin-8 cell adhesion recognition sequence
; US-09-187-859-1699

Query Match          38.3%; Score 23; DB 4; Length 8;
Best Local Similarity 57.1%; Pred. No. 1.9e+05;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 6 DTORFSG 12
   1 :|||||
   1 DTSRFSG 7

Db

RESULT 9
; 5168051-15
; Patent No. 5168051
; APPLICANT: DERYNCK, RIK M.A.; GOEDDEL, DAVID V.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING TGF-B ITS USES
; NUMBER OF SEQUENCES: 21
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/389,929
; FILING DATE: 04-AUG-1989
; SEQ ID NO:15
; LENGTH: 8
; 5168051-15

Query Match          38.3%; Score 23; DB 6; Length 8;
Best Local Similarity 71.4%; Pred. No. 1.9e+05;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 5 LDTORFS 11
   1 :|||
   2 LDTNVS 8

Db

RESULT 10
; US-09-117-860-13
; Sequence 13, Application US/09117860A
; Patent No. 6338955
; GENERAL INFORMATION:
; APPLICANT: OGURI, Suguru
; APPLICANT: MINOWA, Mari
; APPLICANT: YOSHIDA, Aruto
; APPLICANT: TANIGUCHI, Naoyuki
; TITLE OF INVENTION: NOVEL 1-4 N-ACETYLGLUCOSAMINYLTRANSFERASE AND GENE
; FILE REFERENCE: 081356/0119
; CURRENT APPLICATION NUMBER: US/09/117,860A
; CURRENT FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: WO PCT/JP97/04546
; EARLIER FILING DATE: 1997-12-10
; EARLIER APPLICATION NUMBER: JP 161462/1997
; EARLIER FILING DATE: 1997-06-18
; EARLIER APPLICATION NUMBER: JP 332411/1996
; EARLIER FILING DATE: 1996-12-12
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 13
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Bovine
; US-09-117-860-13

Query Match          38.3%; Score 23; DB 4; Length 9;
Best Local Similarity 50.0%; Pred. No. 1.9e+05;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 AEGDTR 9
   1 :|||||
   2 SEGDISK 9

Db

RESULT 11
; US-09-644-600-82
; Sequence 82, Application US/09644600
; Patent No. 6451500
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Tanimoto, Hirotsuhi
; TITLE OF INVENTION: TADG-15: An Extracellular Serine Protease
; FILE REFERENCE: D6064C1P/D
; CURRENT APPLICATION NUMBER: US/09/644,600
; CURRENT FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 09/421,213
; PRIOR FILING DATE: 1999-10-20
; PRIOR APPLICATION NUMBER: 09/027,337
; PRIOR FILING DATE: 1998-02-20
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 82
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Residues 220-228 of the TADG-15 protein
; US-09-644-600-82

Query Match          38.3%; Score 23; DB 4; Length 9;
Best Local Similarity 44.4%; Pred. No. 1.9e+05;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
```

QY 2 AEGIDTOR 10
1 1 1 1
Db 1 ACGINF0K 9

RESULT 12

US-08-599-455B-33
; Sequence 33, Application US/08599455B
; Patent No. 5972621
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; APPLICANT: Tepper, Robert I.
; APPLICANT: Culpeper, Janice A.
; TITLE OF INVENTION: METHODS OF IDENTIFYING COMPOUNDS THAT
; TITLE OF INVENTION: MODULATE BODY WEIGHT USING THE OB RECEPTOR
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: Windows95
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/599,455B
; FILING DATE: 22-JAN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/583,153
; FILING DATE: 28-DEC-1995
; APPLICATION NUMBER: 08/570,142
; FILING DATE: 11-DEC-1995
; APPLICATION NUMBER: 08/569,485
; FILING DATE: 08-DEC-1995
; APPLICATION NUMBER: 08/566,622
; FILING DATE: 04-DEC-1995
; APPLICATION NUMBER: 08/562,663
; FILING DATE: 27-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Melkielejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 07334/017001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-599-455B-33

Query Match 36.7%; Score 22; DB 2; Length 8;
Best Local Similarity 50.0%; Pred. No. 1.9e+05;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 AEGIDTOR 9
1 1 1 1
Db 1 ACGINF0K 8

RESULT 13
US-09-069-781B-33
; Sequence 33, Application US/09069781B
; Patent No. 6287782
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.

; APPLICANT: Tepper, Robert I.
; APPLICANT: Culpeper, Janice A.
; APPLICANT: White, David W.
; TITLE OF INVENTION: THE OB RECEPTOR AND METHODS FOR
; TITLE OF INVENTION: THE DIAGNOSIS AND TREATMENT OF BODY WEIGHT DISORDERS,
; TITLE OF INVENTION: INCLUDING OBESITY AND CACHEXIA
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: Windows95
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/069,781B
; FILING DATE: 29-APRIL-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/864,564
; FILING DATE: 28-MAY-1997
; APPLICATION NUMBER: US 08/708,123
; FILING DATE: 03-SEP-1996
; APPLICATION NUMBER: US 08/638,524
; FILING DATE: 26-APR-1996
; APPLICATION NUMBER: US 08/599,455
; FILING DATE: 22-JAN-1996
; APPLICATION NUMBER: US 08/583,153
; FILING DATE: 28-DEC-1995
; APPLICATION NUMBER: US 08/570,142
; FILING DATE: 11-DEC-1995
; APPLICATION NUMBER: US 08/569,485
; FILING DATE: 08-DEC-1995
; APPLICATION NUMBER: US 08/566,622
; FILING DATE: 04-DEC-1995
; APPLICATION NUMBER: US 08/562,663
; FILING DATE: 27-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Melkielejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 07334/082001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-09-069-781B-33

Query Match 36.7%; Score 22; DB 4; Length 8;
Best Local Similarity 50.0%; Pred. No. 1.9e+05;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 AEGIDTOR 9
1 1 1 1
Db 1 ACGINF0K 8

RESULT 14
US-09-137-132-33
; Sequence 33, Application US/09137132
; Patent No. 6380363
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; APPLICANT: Tepper, Robert I.

APPLICANT: Culpepper, Janice A.
TITLE OF INVENTION: THE OB RECEPTOR AND METHODS FOR
TITLE OF INVENTION: THE DIAGNOSIS AND TREATMENT OF BODY WEIGHT DISORDERS,
TITLE OF INVENTION: INCLUDING OBESITY AND CACHEXIA
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson, P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/137,132
FILING DATE: 18-AUG-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/864,564
FILING DATE: 28-MAY-1997
APPLICATION NUMBER: 08/708,123
FILING DATE: 03-SEP-1996
APPLICATION NUMBER: 08/638,524
FILING DATE: 26-APR-1996
APPLICATION NUMBER: 08/599,455
FILING DATE: 22-JAN-1996
APPLICATION NUMBER: 08/583,153
FILING DATE: 28-DEC-1995
APPLICATION NUMBER: 08/570,142
FILING DATE: 11-DEC-1995
APPLICATION NUMBER: 08/569,485
FILING DATE: 08-DEC-1995
APPLICATION NUMBER: 08/566,622
FILING DATE: 04-DEC-1995
APPLICATION NUMBER: 08/562,663
FILING DATE: 27-NOV-1995
ATTORNEY/AGENT INFORMATION:
NAME: Melkielejohn, Ph.D., Anita L.
REGISTRATION NUMBER: 35,283
REFERENCE/DOCKET NUMBER: 07334/019004
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-5070
TELEFAX: 617-542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-09-137-132-33

Query Match 36.7%; Score 22; DB 4; Length 8;
Best Local Similarity 50.0%; Pred. No. 1.9e+05;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 2 AEGDTR 9
1:1:1:1:
Db 1 AAGLNF 8

RESULT 15
US-08-864-564A-33
Sequence 33, Application US/08864564A
Patent No. 6395498
GENERAL INFORMATION:
APPLICANT: Tartaglia, Louis A.
APPLICANT: Tepper, Robert I.
APPLICANT: Culpepper, Janice A.

APPLICANT: White, David W.
TITLE OF INVENTION: THE OB RECEPTOR AND METHODS FOR
TITLE OF INVENTION: THE DIAGNOSIS AND TREATMENT OF BODY WEIGHT DISORDERS,
TITLE OF INVENTION: INCLUDING OBESITY AND CACHEXIA
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson, P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/864,564A
FILING DATE: 28-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/708,123
FILING DATE: 03-SEP-1996
APPLICATION NUMBER: 08/638,524
FILING DATE: 26-APR-1996
APPLICATION NUMBER: 08/599,455
FILING DATE: 22-JAN-1996
APPLICATION NUMBER: 08/583,153
FILING DATE: 28-DEC-1995
APPLICATION NUMBER: 08/570,142
FILING DATE: 11-DEC-1995
APPLICATION NUMBER: 08/569,485
FILING DATE: 08-DEC-1995
APPLICATION NUMBER: 08/566,622
FILING DATE: 04-DEC-1995
APPLICATION NUMBER: 08/562,663
FILING DATE: 27-NOV-1995
ATTORNEY/AGENT INFORMATION:
NAME: Melkielejohn, Ph.D., Anita L.
REGISTRATION NUMBER: 35,283
REFERENCE/DOCKET NUMBER: 07334/019002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-5070
TELEFAX: 617-542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-864-564A-33

Query Match 36.7%; Score 22; DB 4; Length 8;
Best Local Similarity 50.0%; Pred. No. 1.9e+05;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 2 AEGDTR 9
1:1:1:1:
Db 1 AAGLNF 8

Search completed: April 15, 2003, 10:21:40
job time : 16 secs


```
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 AAEGLDTQRFSG 12
   1:111 1111
Db 123 ADDGLDELRFSG 134

RESULT 5
C86091
Chlamn biosynthesis, pyrindlne moiety [imported] - Escherichia coli (strain O157:H7, s
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: C86091
R:Perma, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Miller, L.; Grobbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamoussis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:1074935; PMID:11206551
A:Accession: C86091
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-631 <STO>
A:Cross-references: GB:AE005174; NID:912518923; PIDN:AAG59191.1; GSPDB:GN00145; UMG:255
C:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: thic
C:Superfamily: thiamin biosynthesis protein thic

Query Match 63.3%; Score 38; DB 2; Length 631;
Best Local Similarity 66.7%; Pred. No. 31;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 AAEGLDTQRFSG 12
   1:111 1111
Db 123 ADDGLDELRFSG 134

RESULT 6
H82574
Phage-related protein XE2291 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: H82574
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: H82574
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-252 <STM>
A:Cross-references: GB:AE004041; GB:AE003849; NID:9107453; PIDN:AAF85090.1; GSPDB:GN001
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; B
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carre
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, B
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawa
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv
M.; Tsubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328
A:Molecule type: DNA
A:Status: preliminary
C:Genetics:
A:Gene: XE2291

Query Match 61.7%; Score 37; DB 2; Length 252;
Best Local Similarity 70.0%; Pred. No. 20;
```

```
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 AAEGLDTQRF 10
   1111 111
Db 240 AAEGLDTQRF 249

RESULT 7
C82548
Phage-related DNA polymerase XF2525 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000
C:Accession: C82548
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: C82548
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-726 <STM>
A:Cross-references: GB:AE004059; GB:AE003849; NID:9107718; PIDN:AAF85323.1; GSPDB:GN
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carre
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; La
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri,
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawa
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv
M.; Tsubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF2525
C:Superfamily: phage SP02 DNA-directed DNA polymerase

Query Match 61.7%; Score 37; DB 2; Length 726;
Best Local Similarity 70.0%; Pred. No. 55;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 AAEGLDTQRF 10
   1111 111
Db 714 AAEGLDTQRF 723

RESULT 8
H82774
Phage-related DNA polymerase XF0683 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000
C:Accession: H82774
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: H82774
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-726 <STM>
A:Cross-references: GB:AE003912; GB:AE003849; NID:9105560; PIDN:AAF83493.1; GSPDB:GN
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carre
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr
```

J.D.; Junqueira, M.L.; Kemper, E.L.; Kitejima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Martino, C.L.; Marques, M.V.; Martins, B
A.Autorio, Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A.Autorio, da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Teshko, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF0683
C:Superfamily: phage SPO2 DNA-directed DNA polymerase

Query Match 61.7%; Score 37; DB 2; Length 726;
Best Local Similarity 70.0%; Pred. No. 53;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 AAGGLDTORF 10
| | | | |
DB 714 AAGGFTYRY 723

RESULT 9

A46546
Leukocyte common antigen long splice form precursor - human
N:Alternate names: CD45; protein-tyrosine-phosphatase, receptor type c; T200 glycoprotein
N:Contains: leukocyte common antigen intermediate splice form; leukocyte common antigen
C:Species: Homo sapiens (man)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jan-2000
C:Accession: A46546; B46546; A29449; B29449; I57658
R:Stratell, M.; Hall, L.R.; Sagar, Y.; Schlossman, S.F.; Saito, H.
J. Exp. Med. 166, 1548-1566, 1987
A>Title: Differential usage of three exons generates at least five different mRNAs encod
A:Reference number: A46546; MUID:88061067; PMID:2824653
A:Accession: A46546
A:Molecule type: mRNA
A:Status: preliminary
A:Residues: 1-32,99-264 <STR>
A:Cross-references: GB:Y00638
A:Experimental source: clone LCA.111 and clone LCA.260
A:Accession: C46546
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-32,99-264 <STR>
A:Cross-references: GB:Y00638
A:Experimental source: clone LCA.111 and clone LCA.260
A:Accession: C46546
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-31,193-264 <STR>
A:Cross-references: GB:Y00638
A:Experimental source: clone LCA.1
R:Halp, S.J.; Thomas, M.L.; Morton, C.C.; Trowbridge, I.S.
EMBO J. 6, 1251-1257, 1987
A>Title: Structural variants of human T200 glycoprotein (leukocyte-common antigen).
A:Reference number: A91066; MUID:87275816; PMID:2996090
A:Accession: A29449
A:Molecule type: mRNA
A:Residues: 1-31,193-649, 'L', 651-869, 'G', 871-872, 'A', 874-1206, 'P', 1208-1304 <RAL>
A:Cross-references: GB:Y00062; NID:934275; PIDN:CAA68269.1; PID:934276
A:Experimental source: clones pHC-1 and lambdaHLG1
A:Accession: B29449
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 32-192 <RA2>
A:Experimental source: clone HLC-2
R:Tsal, A.Y.; Streuli, M.; Saito, H.
Mol. Cell. Biol. 9, 4550-4555, 1989
A>Title: Integrity of the exon 6 sequence is essential for tissue-specific alternative s
A:Reference number: I57658; MUID:90066468; PMID:2531281
A:Accession: I57658
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA

A:Residues: 146-192 <RES>
A:Cross-references: GB:I29253; NID:9187020; PIDN:AA59497.1; PID:9553521
C:Genetics:
A:Gene: GDB:PPRC; CD45
A:Cross-references: GDB:119768; OMIM:151460
A:Map position: 1931-1932
C:Superfamily: leukocyte common antigen; leukocyte common antigen cytosolic domain ho
C:Keywords: alternative splicing; glycoprotein; phosphoprotein; phosphoric monoester
F:594-1235/Domain: leukocyte common antigen cytosolic domain homology <LAC>
F:675-899/Domain: protein-tyrosine-phosphatase homology <PPP>
F:851/Active site: Cys (phosphotyrosine intermediate) #status predicted
F:857/Binding site: substrate phosphate (Arg) #status predicted

Query Match 60.0%; Score 36; DB 1; Length 1304;
Best Local Similarity 77.8%; Pred. No. 15e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 GLDPTQRFSG 12
| | | | |
DB 119 GTDPTQRFSG 127

RESULT 10

S25663
T-cell surface glycoprotein CD8 alpha chain - bovine
C:Species: Bos primigenius taurus (cattle)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 23-Jul-1999
C:Accession: S25663
R:Alator, P.; Bucet, C.; Fornaro, M.; Rattazzi, M.C.; Nakauchi, H.; Herzenberg, L.A.;
Immunology 76, 95-102, 1992
A>Title: Molecular cloning, reconstruction and expression of the gene encoding the al
A:Reference number: S25663; MUID:92332098; PMID:1628904
A:Accession: S25663
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-242 <TAI>
A:Cross-references: EMBL:X59416; NID:9190; PIDN:CAA42051.1; PID:9191
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: glycoprotein; transmembrane protein

Query Match 58.3%; Score 35; DB 2; Length 242;
Best Local Similarity 63.6%; Pred. No. 46;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 AEGDPTQRFSG 12
| | | | |
DB 85 AEGDPTQRFSG 95

RESULT 11

AE0154
Probable nucleoid-associated protein YPO1262 [imported] - Yersinia pestis (strain CO9
C:Species: Yersinia pestis
C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
C:Accession: AE0154
R:Parikh, J.; When, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, M
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G
ll, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barre
Nature 413, 523-527, 2001
A>Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AE0154
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-334 <KUR>
A:Cross-references: GB:AL590842; PIDN:CA90096.1; PID:915979316; GSPDB:GN00175
C:Genetics:
A:Gene: YPO1262

Query Match 58.3%; Score 35; DB 2; Length 334;
Best Local Similarity 87.5%; Pred. No. 63;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAEGLDTRQ 8
 |||||
 DB 188 AAEGLDTRK 195

RESULT 12

F95953
 probable glycosyltransferase, forming alpha-glycosyl linkages protein expc [imported] -
 C:Species: *Sinorhizobium meliloti*
 C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
 C:Accession: F95953
 R:Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Hernan
 Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
 A:Title: The complete sequence of the 1,583-kb pSymb megaplasmid from the N2-fixing endo
 A:Reference number: A95842; MUID:21396508; PMID:11481431
 A:Accession: F95953
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-417 <KUR>
 A:Cross-references: GB:AL591985; PIDN:CAC49294.1; PID:915140780; GSPDB:GN00167
 A:Experimental source: strain 1021, megaplasmid pSymb
 R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
 Pella, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
 L.; Hyman, R.W.; Jones, T.
 Science 293, 668-672, 2001
 A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
 hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
 A:Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.
 A:Reference number: A96039; MUID:21368234; PMID:11474104
 A:Contents: annotation
 C:Genetics:
 A:Gene: expc: SMB21318
 A:Genome: plasmid

Query Match 58.3%; Score 35; DB 2; Length 417;
 Best Local Similarity 75.0%; Pred. No. 78;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 3 EGLDTRQ 10
 |||||
 DB 188 EGVDTTRF 195

RESULT 13

C81965
 hypothetical periplasmic protein NMA0476 [imported] - *Neisseria meningitidis* (strain Z24
 C:Species: *Neisseria meningitidis*
 C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
 C:Accession: C81965
 R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel,
 ; Holroyd, S.; Jørgensen, K.; Leather, S.; Moutle, S.; Mungall, K.; Quail, M.A.; Rajandream,
 Nature 404, 502-506, 2000
 A:Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491.
 A:Reference number: AB1775; MUID:20222556; PMID:10761919
 A:Accession: C81965
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-505 <PAR>
 A:Cross-references: GB:AL162753; GB:AL157959; NID:97379120; PIDN:CAB83771.1; PID:9737922
 A:Experimental source: serogroup A, strain Z2491
 C:Genetics:
 A:Gene: NMA0476

Query Match 58.3%; Score 35; DB 2; Length 505;
 Best Local Similarity 72.7%; Pred. No. 94;
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 AAEGLDTRFSG 12
 |||||
 DB 184 AAEGLDTRFSG 194

RESULT 14

F87664
 glycine cleavage system P protein, subunit 2 [imported] - *Caulobacter crescentus*
 C:Species: *Caulobacter crescentus*
 C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
 C:Accession: F87664
 R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Ko
 n. J.; Esmolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
 A:Title: Complete Genome Sequence of *Caulobacter crescentus*.
 A:Reference number: AB7249; MUID:21173698; PMID:11259647
 A:Accession: F87664
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-524 <SNO>
 A:Cross-references: GB:AE005673; NID:913425054; PIDN:AAK25314.1; GSPDB:GN00148
 C:Genetics:
 A:Gene: CC3352

Query Match 58.3%; Score 35; DB 2; Length 524;
 Best Local Similarity 58.3%; Pred. No. 97;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 AAEGLDTRFSG 12
 |||||
 DB 479 AAKAGDTRFSG 490

RESULT 15

E70752
 probable dead protein - *Mycobacterium tuberculosis* (strain H37RV)
 C:Species: *Mycobacterium tuberculosis*
 C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 02-Feb-2001
 C:Accession: E70752
 R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
 ; Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holroyd,
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete geno
 A:Reference number: A70500; MUID:96295987; PMID:9634230
 A:Accession: E70752
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-563 <COL>
 A:Cross-references: GB:277137; GB:AL123456; NID:93261593; PIDN:CAB00899.1; PID:e25488
 A:Experimental source: strain H37RV
 C:Genetics:
 A:Gene: dead
 C:Keywords: ATP; nucleotide binding; P-loop
 F:57-64/Region: nucleotide-binding motif A (P-loop)
 F:159-164/Region: nucleotide-binding motif B
 F:163-166/Region: DEAD motif

Query Match 58.3%; Score 35; DB 2; Length 563;
 Best Local Similarity 63.6%; Pred. No. 1e+02;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 AAEGLDTRFSG 11
 |||||
 DB 310 AAEGLDTRFSG 320

Search completed: April 15, 2003, 10:14:47
 Job time : 45 secs

GenCore version 5.1.4_p5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 15, 2003, 10:11:05 ; Search time 11 Seconds

(without alignments)
45.247 Million cell updates/sec

Title: US-09-521-527C-1

Sequence: 1 AAEGDTRFSG 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	60	100.0	198	1	CD8A_PONPY
2	60	100.0	235	1	CD8A_HUMAN
3	41	68.3	239	1	CD8A_CANFA
4	38	63.3	631	1	THIC_ECO57
5	38	63.3	631	1	THIC_ECOLI
6	36	60.0	1304	1	CD45_HUMAN
7	35	58.3	242	1	CD8A_BOVIN
8	35	58.3	563	1	DEAD_MYCTU
9	35	58.3	628	1	DEAD_ECOLI
10	35	58.3	631	1	THIC_SALTI
11	35	58.3	631	1	THIC_SALTY
12	35	58.3	642	1	DEAD_KLEPN
13	35	58.3	829	1	GUTR_BACSU
14	35	58.3	1134	1	PHY1_SELMA
15	35	58.3	1515	1	GUTR_AZORA
16	34.5	57.5	290	1	AMPR_CITFR
17	34.5	57.5	290	1	AMPR_ENTCL
18	34	56.7	90	1	YGGX_HAEIN
19	34	56.7	203	1	AROG_AMEME
20	34	56.7	375	1	INTE_ECOLI
21	34	56.7	380	1	VINT_BPR21
22	34	56.7	552	1	CHOD_BREST
23	34	56.7	693	1	VG50_BPMD2
24	34	56.7	849	1	PHSG_SYNY3
25	34	56.7	1056	1	DPOL_ADE02
26	34	56.7	1056	1	DPOL_ADE05
27	33.5	55.8	706	1	CATE_MYCAV
28	33	55.0	328	1	AFUC_HAEIN
29	33	55.0	372	1	YEF1_ECOLI
30	33	55.0	432	1	DADI_PSEAE
31	33	55.0	484	1	NTRC_RHIME
32	33	55.0	527	1	PTB_MOUSE
33	33	55.0	531	1	PTB_HUMAN

34	33	55.0	555	1	PTB_RAT	000438	rattus norv
35	33	55.0	601	1	DEAD_BUCAI	P57453	buchnera ap
36	33	55.0	714	1	YEF7_YEAST	P43556	saccharomyc
37	33	55.0	948	1	RCSC_SALTI	056128	salmonella
38	33	55.0	948	1	RCSC_SALTY	P58662	salmonella
39	33	55.0	1193	1	DPOL_ADE04	P87503	human adeno
40	33	55.0	1325	1	YAE6_SCHPO	009847	schizosacch
41	33	55.0	2132	1	PGCA_MOUSE	061282	mus musculu
42	32.5	54.2	296	1	AMPR_PSEAE	P24734	pseudomonas
43	32.5	54.2	930	1	SM6C_HUMAN	Q9h3t2	homo sapien
44	32.5	54.2	931	1	SM6C_MOUSE	Q9wtm3	mus musculu
45	32.5	54.2	960	1	SM6C_RAT	Q9wtl3	rattus norv

ALIGNMENTS

RESULT 1	CD8A_PONPY	STANDARD;	PRT;	198 AA.
ID	CD8A_PONPY			
AC	P30433;			
DT	01-APR-1993 (Rel. 25, Created)			
DT	01-APR-1993 (Rel. 25, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	T-cell surface glycoprotein CD8 alpha chain precursor (T-lymphocyte			
DE	differentiation antigen T8/LEU-2).			
GN	CD8A.			
OS	Pongo pygmaeus (Orangutan).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pongo.			
OX	NCBI_TaxID=9600;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Isolate Jari;			
RX	MEDLINE=92307742; PubMed=1612644;			
RA	Lawlor D.A., Parham P.;			
RT	"Structure of CD8 alpha and beta chains of the orangutan: novel			
RT	patterns of mRNA splicing encoding hingeless polypeptides."			
RL	Immunogenetics 36:121-125(1992).			
CC	- FUNCTION: IDENTIFIES CYTOTOXIC/SUPPRESSOR T-CELLS THAT INTERACT			
CC	WITH MHC CLASS I BEARING TARGETS. CD8 IS THOUGHT TO PLAY A ROLE IN			
CC	THE PROCESS OF T-CELL MEDIATED KILLING. CD8 ALPHA CHAINS BINDS TO			
CC	CLASS I MHC MOLECULES ALPHA-3 DOMAINS.			
CC	- SUBUNIT: IN GENERAL HETERODIMER OF AN ALPHA AND A BETA CHAIN			
CC	LINKED BY TWO DISULFIDE BONDS. CAN ALSO FORM HOMODIMERS.			
CC	- SUBCELLULAR LOCATION: Type I membrane protein.			
CC	- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.			
CC	- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.			
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	EMBL: X60223; CAA42784.1; -			
DR	HSSP: P01732; ICD8			
DR	InterPro: IPR003006; Ig_MHC.			
DR	InterPro: IPR003596; Ig_V.			
DR	Pfam: PF00047; Ig_1.			
DR	SMART: SM00406; IGV_1.			
KW	Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell;			
KW	Immune response; Signal.			
FT	SIGNAL	1	21	
FT	CHAIN	22	198	
FT				
FT	DOMAIN	22	145	
FT	TRANSMEM	146	166	
FT	DOMAIN	167	198	
FT	DOMAIN	22	135	
FT	DOMAIN	43	115	
FT	DISULFID			
				BY SIMILARITY.
				T-CELL SURFACE GLYCOPROTEIN CD8 ALPHA
				CHAIN.
				EXTRACELLULAR (POTENTIAL).
				POTENTIAL.
				CYTOPLASMIC (POTENTIAL).
				IG-LIKE V-TYPE DOMAIN.
				BY SIMILARITY.


```

DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Thiamine biosynthesis protein thtC.
GN thtC OR Z5569 OR ECS4917.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G., III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Postal G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Gottlieb E.J., Davis N.W., Lam A., Dimalanta E.T., Potamoudis K.,
RA Alachkar A., Anantharaman T.S., Iln J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
RL Nature 409:529-533(2001).
RN [2]
RN [2]
RC SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakiwa C., Ogasawara N., Yasunaga T.,
RA Kohara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
O157:H7 and genomic comparison with a laboratory strain K-12.";
DL DNA Res. 8:11-22(2001).
CC -1- FUNCTION: Required for the synthesis of the hydromethylpyrimidine
CC (HMP) moiety of thiamine (4-amino-2-methyl-5-
CC hydroxymethylpyrimidine) (By similarity).
CC -1- PATHWAY: Thiamine biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE thtC FAMILY.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its use
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE005631; AAC59191.1; -
DR EMBL: AP002567; BAB38340.1; -.
DR InterPro: IPR002817; thtC.
DR Pfam: PF01964; thtC; 1.
DR ProDom: PD007048; thtC; 1.
DR TRGFAMS: TIGR00190; thtC; 1.
KW Thiamine biosynthesis; Complete proteome.
SQ
SEQUENCE 631 AA; 70614 MW; 5494CBBD9C036F2B CRC64;

Query Match 63.3%; Score 38; DB 1; Length 631;
Best Local Similarity 66.7%; Pred. No. 7.5;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 AAEGLDPTQRESG 12
   1 : ||| |||
Db 123 ADDGLDELREFG 134

RESULT 5
THtC_ECOLI STANDARD; PRT; 631 AA.
AC P30136;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Thiamine biosynthesis protein thtC.

```

OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12;
 RX MEDLINE=93163063; PubMed=8432721;
 RA Vander Horn P.B., Backstrom A.D., Stewart V., Begley T.P.;
 RT "Structural genes for thiamine biosynthetic enzymes (thiCEGH) in
 RT Escherichia coli K-12.";
 RL J. Bacteriol. 175:982-992(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655;
 RX MEDLINE=94089392; PubMed=8265357;
 RA Blatterer F.R., Burland V.D., Plunkett G. III, Sofia H.J.,
 RA Daniels D.L.;
 RT "Analysis of the Escherichia coli genome. IV. DNA sequence of the
 RT region from 89.2 to 92.8 minutes.";
 RL Nucleic Acids Res. 21:5408-5417(1993).
 CC -1- FUNCTION: REQUIRED FOR THE SYNTHESIS OF THE HYDROMETHYLPYRIMIDINE
 CC (HMP) MOIETY OF THIAMINE (4-AMINO-2-METHYL-5-
 CC HYDROXYMETHYLPYRIMIDINE).
 CC -1- PATHWAY: Thiamine biosynthesis.
 CC -1- SIMILARITY: BELONGS TO THE THIC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M88701; AAB95616.1; -;
 DR EMBL: U00006; AAC43092.1; -;
 DR EMBL: AE000473; AAC76968.1; -;
 DR PIR: S53117; S53117.
 DR Ecogene: EGI1585; thic.
 DR InterPro: IPR002817; Thic.
 DR Pfam: PF01964; thic; 1.
 DR ProDom: PD007048; thic; 1.
 DR TIGRfam: TIGR00190; thic; 1.
 KW Thiamine biosynthesis; Complete proteome.
 FT CONFLICT 623 630 EYLRKEE -> RNL P (IN REF. 1).
 SQ SEQUENCE 631 AA; 70850 MW; 6P2105B4792CBA3F CRC64;
 Query Match 63.3%; Score 38; DB 1; Length 631;
 Best Local Similarity 66.7%; Pred. No. 7.5;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 OY 1 AAEGLDTRFGS 12
 DB 123 ADDGLDELRFSG 134
 ID CD45_HUMAN STANDARD; PRT; 1304 AA.
 AC P08575;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Leukocyte common antigen precursor (EC 3.1.3.48) (L-CA) (CD45 antigen)
 DE (T200)
 GN PTPRC OR CD45.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.

RC TISSUE=Lymphocytes;
 RX MEDLINE=88061067; PubMed=2824653;
 RA Streuli M., Hall L.R., Saga Y., Schlossman S.F., Saito H.;
 RT "Differential usage of three exons generates at least five different
 RT mRNAs encoding human leukocyte common antigens.";
 RL J. Exp. Med. 166:1548-1566(1987).
 RN [2]
 RP FUNCTION
 RX MEDLINE=89017162; PubMed=2845400;
 RA Charbonneau H., Tonks N.K., Walsh K.A., Fischer E.H.;
 RT "The leukocyte common antigen (CD45): a putative receptor-linked
 RT protein tyrosine phosphatase.";
 RL Proc. Natl. Acad. Sci. U.S.A. 85:7182-7186(1988).
 RN [3]
 RP MYOGENESIS.
 RX MEDLINE=90316093; PubMed=1695146;
 RA Streuli M., Krieger N.X., Thai T., Tang M., Saito H.;
 RT "Distinct functional roles of the two intracellular phosphatase like
 RT domains of the receptor-linked protein tyrosine phosphatases LCA and
 RT LAR.";
 RL EMBL J. 9:2399-2407(1990).
 CC -1- FUNCTION: REQUIRED FOR T-CELLS ACTIVATION THROUGH THE ANTIGEN
 CC RECEPTOR. THE FIRST PTPASE DOMAIN HAS ENZYMAIC ACTIVITY, WHILE
 CC THE SECOND ONE SEEMS TO AFFECT THE SUBSTRATE SPECIFICITY OF THE
 CC FIRST ONE.
 CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
 CC tyrosine + phosphate.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 8 ISOFORMS ARE PRODUCED BY
 CC ALTERNATIVE SPLICING.
 CC -1- PTM: HEAVILY N- AND O-GLYCOSYLATED.
 CC -1- SIMILARITY: CONTAINS 2 PROTEIN-TYROSINE PHOSPHATASE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
 CC -1- DATABASE: NAME=PROM; NOTE=CD guide CD45 entry;
 CC WWW=<http://www.ncbi.nlm.nih.gov/prov/cd/cd45.htm>.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Y00638; CA68669.1; -;
 DR HSSP: P18052; IYFO.
 DR GlycoSiteDB: P08575; -;
 DR Genew: HGNC:9666; PTPRC.
 DR MIM: 151460; -;
 DR InterPro: IPR003961; FN_III.
 DR InterPro: IPR000387; TYR-phosphatase.
 DR InterPro: IPR000242; TYL_PP.
 DR Pfam: PF00041; fn3; 2.
 DR Pfam: PF00102; Y-phosphatase; 2.
 DR PRINTS: PR00700; PRTYPHPTASE.
 DR SMART: SM00060; FN3; 2.
 DR SMART: SM00194; PTPC; 2.
 DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 2.
 DR PROSITE: PS0056; TYR_PHOSPHATASE_2; 2.
 DR PROSITE: PS50055; TYR_PHOSPHATASE_PTP; 2.
 KW Glycoprotein; Transmembrane; Phosphorylation; B-cell; T-cell; Repeat;
 KW Alternative splicing; Hydrolase; Signal.
 FT SIGNAL 1 23
 FT CHAIN 24 1304
 FT DOMAIN 24 575
 FT TRANSMEM 576 597
 FT DOMAIN 598 1304
 FT DOMAIN 387 479
 FT DOMAIN 480 571
 FT DOMAIN 670 919
 FT DOMAIN 961 1235
 FT ACT_SITE 851 851
 FT ACT_SITE 1167 1167

```
FT CARBOHYD 78 78 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 90 90 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 95 95 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 190 190 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 232 232 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 260 260 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 270 270 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 335 335 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 378 378 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 419 419 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 468 468 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 488 488 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 529 529 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT MUTAGEN 851 851 C->S: LOSS OF ACTIVITY.
SQ SEQUENCE 1304 AA; 147253 MW; 1F357BC563261882 CRC64;

Query Match
Best Local Similarity 60.0%; Score 36; DB 1; Length 1304;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 GDTDTQFSG 12
DB 119 GDTDTQFSG 127

RESULT 7
CD8A_BOVIN STANDARD; PRT; 242 AA.
AC P31783;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE T-cell surface glycoprotein CD8 alpha chain precursor.
GN CD8A.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxId=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Thymus;
RX MEDLINE=92332098; PubMed=1628904;
RA Lalor P., Buccì C., Fornaro M., Rattazzi M.C., Nakauchi H.,
RA Herzenberg L.A., Alberti S.;
RT "Molecular cloning, reconstruction and expression of the gene
RT encoding the alpha-chain of the bovine CD8 -- definition of three
RT peptide regions conserved across species.";
RL Immunology 76:95-102(1992).
CC -1- FUNCTION: IDENTIFIES CYTOTOXIC/SUPPRESSOR T-CELLS THAT INTERACT
CC WITH MHC CLASS I BEARING TARGETS. CD8 IS THOUGHT TO PLAY A ROLE IN
CC THE PROCESS OF T-CELL MEDIATED KILLING. CD8 ALPHA CHAINS BINDS TO
CC CLASS I MHC MOLECULES ALPHA-3 DOMAINS.
CC -1- SUBUNIT: IN GENERAL, HETERODIMER OF AN ALPHA AND A BETA CHAIN
CC LINKED BY TWO DISULFIDE BONDS. CAN ALSO FORM HOMODIMERS.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X59416; CAA42051.1; -
CC PIR: S25663; S25663.
CC HSSP: P01732; IC08.
```

```
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IgV; 1.
KM Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell;
KW Immune response; Signal.
FT STGNAL 1 25
FT CHAIN 26 242
FT
FT DOMAIN 26 189
FT TRANSMEM 190 214
FT DOMAIN 215 242
FT DOMAIN 26 140
FT DISULFID 47 120
SQ SEQUENCE 242 AA; 26417 MW; 91481320EF05195E CRC64;

Query Match
Best Local Similarity 58.3%; Score 35; DB 1; Length 242;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 AEGIDTQFSG 12
DB 85 AEGIDTQFSG 95

RESULT 8
DEAD_MYCTU STANDARD; PRT; 563 AA.
AC Q11039;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cold-shock DEAD-box protein A homolog (Amp-dependent RNA helicase dead
DE homology).
GN DEAD OR RV1253 OR MT1292 OR MYC50.29C.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxId=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feldwell T., Gentles S., Hamlin N., Holtroyd S.,
RA Hornby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skellon S., Squares R.,
RA Sultson J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson S., DeBoy R., Dodson R., Gwin M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikala A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: HAS A HELIX-DESTABILIZING ACTIVITY (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
```

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: 277137; CAB00899.1; -
 CC EMBL: AE007004; AAK45550.1; -
 DR HSSP: 058083; 1HV8.
 DR TIGR: MT1292; -
 DR Tuberculist; RV1253; -
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR000629; DEAD_box.
 DR InterPro: IPR001650; Helicase_C.
 DR Pfam: PF00270; DEAD.1.
 DR Pfam: PF00271; Helicase_C.1.
 DR SMART: SM00487; DEXDC.1.
 DR SMART: SM00490; HELICG.1.
 DR PROSITE: PS00039; DEAD_ATP_HELICASE.1.
 KW Hydrolyase: Helicase; ATP-binding; RNA-binding;
 KW Transcription regulation; Complete proteome.
 FT NP_BIND 57 64 ATP (BY SIMILARITY).
 FT SITE 163 166 DEAD BOX.
 SQ SEQUENCE 563 AA: 61452 MW: 5D6D0FAE15D3ADE CRC64;

Query Match 58.3%; Score 35; DB 1; Length 563;
 Best Local Similarity 63.6%; Pred. No. 27;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 AAEGIDTQRF 11
 DB 310 AARGIDVERIS 320

RESULT 9
 ID DEAD_ECOLI STANDARD: PRT; 628 AA.
 AC P23304;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cold-shock DEAD-box protein A (ATP-dependent RNA helicase dead).
 GN DEAD OR CSDA OR MSSB OR B3162.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxId=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91258309; PubMed=2045359;
 RA Toome W.M., Rudd K.E., Friesen J.D.;
 RT "dead, a new Escherichia coli gene encoding a presumed ATP-dependent
 RT RNA helicase, can suppress a mutation in rpsB, the gene encoding
 RT ribosomal protein S2";
 RT J. Bacteriol. 173:3291-3302(1991).
 RL [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9378503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12";
 RL Science 277:1453-1474(1997).
 RN [3]
 RP SEQUENCE OF 518-628 FROM N.A.
 RC STRAIN=JM101;
 RA Peng H., Hsieh M., Zao C., Chang H.-Y.;
 RT "Nucleotide sequence and expression in Escherichia coli of the
 RT Klebsiella pneumoniae dead gene";
 RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 1-20, AND CHARACTERIZATION.
 RX MEDLINE=96133880; PubMed=8552679;

RA Jones P.G., Mita M., Kim Y., Jiang W., Inouye M.;
 RT "Cold shock induces a major ribosomal-associated protein that unwinds
 RT double-stranded RNA in Escherichia coli";
 RL Proc. Natl. Acad. Sci. U.S.A. 93:76-80(1996).
 CC -1- FUNCTION: HAS A HELIX-DESTABILIZING ACTIVITY. PLAYS A KEY ROLE IN
 CC OPTIMAL CELL GROWTH AT LOW TEMPERATURE AND IS REQUIRED FOR NORMAL
 CC CELL DIVISION. SUPPRESSOR OF A MUTANT DEFECTIVE IN RPSB GENE FOR
 CC RIBOSOMAL PROTEIN S2.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -1- INDUCTION: IN RESPONSE TO LOW TEMPERATURE.
 CC -1- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY.
 CC CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN FROM POSITION 535
 CC ONWARD AND IS SHORTER (553 AA) DUE TO A FRAMESHIFT.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: M63288; AAA23674.1; ALT_FRAME.
 CC EMBL: U18997; AAA57965.1; ALT_INIT.
 CC EMBL: AE000397; AAC76196.1; ALT_INIT.
 CC EMBL: U03750; AAA03626.1; -
 CC PIR: A42357; A42357.
 CC HSSP: 058083; 1HV8.
 CC Ecogene; EG10215; dead.
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR000629; DEAD_box.
 DR InterPro: IPR001650; Helicase_C.
 DR Pfam: PF00270; DEAD.1.
 DR Pfam: PF00271; Helicase_C.1.
 DR SMART: SM00487; DEXDC.1.
 DR SMART: SM00490; HELICG.1.
 DR PROSITE: PS00039; DEAD_ATP_HELICASE.1.
 KW Hydrolyase: Helicase; ATP-binding; RNA-binding;
 KW Transcription regulation; Complete proteome.
 FT INIT_MET 0
 FT NP_BIND 49 56 ATP (BY SIMILARITY).
 FT SITE 155 158 DEAD BOX.
 FT DOMAIN 567 628 ARG/GLU/GTY-RICH.
 FT CONFLICT 443 443 G -> A (TN REF.2).
 SQ SEQUENCE 628 AA: 70400 MW: 451528421DA26D99 CRC64;

Query Match 58.3%; Score 35; DB 1; Length 628;
 Best Local Similarity 63.6%; Pred. No. 30;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 AAEGIDTQRF 11
 DB 304 AARGIDVERIS 314

RESULT 10
 ID THIC_SALTI STANDARD: PRT; 631 AA.
 AC Q86326;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Thiamine biosynthesis protein thic.
 GN THIC OR STY3721.
 OS Salmonella typhi.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxId=601;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CT18;
 RX MEDLINE=21534947; PubMed=11677608;
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,

RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahinia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltyell T., Hamlin N., Haque A., Hien T.T., Holtroyd S., Jagels K.,
RA Krogsh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrall B.G.,
RT Complete genome sequence of a multiple drug resistant *Salmonella*
RT *enterica* serovar Typh1 CT18.*;
RL Nature 413:848-852(2001).
CC -1- FUNCTION: Required for the synthesis of the hydroxymethylpyrimidine
CC (HMP) moiety of thiamine (4-amino-2-methyl-5-
CC hydroxymethylpyrimidine) (By similarity).
CC -1- PATHWAY: Thiamine biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE THIC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AL627279; CAD09480.1; -
CC InterPro: IPR002817; Thic.
CC Pfam: PF01964; Thic; 1.
CC ProDom: PD007048; Thic; 1.
CC TIGRFAMs: TIGR00190; thic; 1.
CC Thiamine biosynthesis; Complete proteome.
CC KW Thiamine biosynthesis; Complete proteome.
CC SEQUENCE 631 AA; 70802 MW; A323C630338EFAAD CRC64;
SO
Query Match 58.3%; Score 35; DB 1; Length 631;
Best Local Similarity 58.3%; Pred. NO. 30;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
OY 1 AAEGLDTRFSG 12
Db 123 ADDGLDLRFTG 134
RESULT 11
THIC_SALTY
ID THIC_SALTY STANDARD; PRT; 631 AA.
AC Q9L917;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Thiamine biosynthesis protein thic.
GN THIC OR STM4164 OR STM1.33.
OS *Salmonella typhimurium*.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC *Salmonella*.
OX NCBI_TaxID=602;
ON [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGGC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Muiyaney E.,
RA Ryan E., Sun H., Florita L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT Complete genome sequence of *Salmonella enterica* serovar Typhimurium
RT LT2.*;
RL Nature 413:852-856(2001).
CC -1- FUNCTION: REQUIRED FOR THE SYNTHESIS OF THE HYDROMETHYLPYRIMIDINE
CC (HMP) MOIETY OF THIAMINE (4-AMINO-2-METHYL-5-
CC HYDROXYMETHYLPYRIMIDINE) (BY SIMILARITY).
CC -1- PATHWAY: Thiamine biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE THIC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF170176; AAF33523.1; -
CC DR EMBL: AE008894; ALU22992.1; -
CC DR StyGene: SG27272; thic.
CC DR InterPro: IPR002817; Thic.
CC Pfam: PF01964; Thic; 1.
CC ProDom: PD007048; Thic; 1.
CC TIGRFAMs: TIGR00190; thic; 1.
CC Thiamine biosynthesis; Complete proteome.
CC KW Thiamine biosynthesis; Complete proteome.
CC SEQUENCE 631 AA; 70845 MW; 0A298124EC6644A CRC64;
SO
Query Match 58.3%; Score 35; DB 1; Length 631;
Best Local Similarity 58.3%; Pred. NO. 30;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
OY 1 AAEGLDTRFSG 12
Db 123 ADDGLDLRFTG 134
RESULT 12
DEAD_KLEPN
ID DEAD_KLEPN STANDARD; PRT; 642 AA.
AC P33906;
DT 01-FEB-1994 (Rel. 28, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cold-shock DEAD-box protein A (ATP-dependent RNA helicase dead).
GN DEAD OR CSDA.
OS *Klebsiella pneumoniae*.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC *Klebsiella*.
OX NCBI_TaxID=573;
ON [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CG43;
RX Peng H.L., Hsieh M.-J., Zao C.-L., Chang H.-Y.;
RA "Nucleotide sequence and expression in *Escherichia coli* of the
RT *Klebsiella pneumoniae* dead gene.";
RL J. Biochem. 115:409-414(1994).
CC -1- FUNCTION: HAS A HELIX-DESTABILIZING ACTIVITY. PLAYS A KEY ROLE IN
CC OPTIMAL CELL GROWTH AT LOW TEMPERATURE AND IS REQUIRED FOR NORMAL
CC CELL DIVISION (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L08387; AAA61345.1; ALT_INIT.
CC DR PIR: JX0314; JX0314.
CC DR HSSP: Q58083; 1HW.
CC DR InterPro: IPR001410; DEAD.
CC DR InterPro: IPR000629; DEAD-box.
CC DR InterPro: IPR001650; Helicase_C.
CC Pfam: PF00270; DEAD; 1.
CC Pfam: PF00271; Helicase_C; 1.
CC SMART: SM00487; DEXDC; 1.
CC DR SMART: SM00480; HELICG; 1.
CC PROSITE: PS00039; DEAD_ATP_HELICASE; 1.
CC Hydrolyase; Helicase; ATP-binding; RNA-binding;


```

CC -1- SUBUNIT: HOMODIMER.
CC -1- PTM: CONTAINS ONE COVALENTLY LINKED TETRACYCLOLE CHROMOPHORE.
CC -1- SIMILARITY: BELONGS TO THE PHYTOCHROME FAMILY.
CC -1- SIMILARITY: CONTAINS 2 PAS (PER-ARNT-SIM) DIMERIZATION DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 PAS-ASSOCIATED C-TERMINAL (PAC) DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 HISTIDINE KINASE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X61458; CAA43698.1; -.
CC PIR: S31280; S31280.
CC InterPro: IPR003594; ATPbind_ATPase.
CC InterPro: IPR003018; GAF.
CC InterPro: IPR004359; HIS_KIN_sig.
CC InterPro: IPR003661; HIS_KIN.
CC InterPro: IPR001610; PAC.
CC InterPro: IPR000700; PAS-assoc.C.
CC InterPro: IPR001294; Phytochrome.
CC InterPro: IPR001294; Phytochrome.
CC Pfam: PF00360; Phytochrome; 1.
CC Pfam: PF00512; signal; 1.
CC Pfam: PF00989; PAS; 2.
CC Pfam: PF01590; GAF; 1.
CC Pfam: PF02518; HATPase_C; 1.
CC PRINTS: PRO1033; PHYTOCHROME.
CC SMART: SM00065; GAF; 1.
CC SMART: SM00387; HATPase_C; 1.
CC SMART: SM00388; HSKA; 1.
CC SMART: SM00086; PAC; 1.
CC SMART: SM00091; PAS; 2.
CC TIGRfam: TIGR00229; sensory_box; 2.
CC PROSITE: PS50109; HIS_KIN; 1.
CC PROSITE: PS50113; PAC; 2.
CC PROSITE: PS50112; PAS; 2.
CC PROSITE: PS50245; PHYTOCHROME_1; 1.
CC PROSITE: PS50245; PHYTOCHROME_2; 1.
CC Transcription regulation: Photoreceptor: Phytochrome; Chromophore;
CC Repeat; Multigene family.
CC KW DOMAIN 616 687 PAS 1.
CC FT DOMAIN 690 746 PAC.
CC FT DOMAIN 750 821 PAS 2.
CC FT DOMAIN 901 1121 HISTIDINE KINASE.
CC FT BINDING 324 324 CHROMOPHORE (BY SIMILARITY).
CC FT DOMAIN 352 358 POLY-GLY.
CC SEQENCE 1134 AA; 124706 MW; 42819B9FAAC398C CRC64;

Query Match 58.3%; Score 35; DB 1; Length 1134;
Best Local Similarity 60.0%; Pred. NO. 57;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 AEGLDTORF 10
   11:111111
Db 827 AADGQDETKF 836

RESULT 15
GLTB.AZOBR STANDARD; PRT; 1515 AA.
ID GLTB.AZOBR
AC 005755;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glutamate synthase [NADPH] large chain precursor (EC 1.4.1.13)
DE (Glutamate synthase alpha subunit) (NADPH-GOGAT) (GLTs alpha chain).
GN GLTB.
OS Azospirillum brasilense.
OC Bacteria; Proteobacteria; alpha subdivision; Rhodospirillaceae;

```

```

CC Azospirillum.
CC NCBI_Taxid=192;
CC [1]
CC SEQUENCE FROM N.A., AND SEQUENCE OF 37-56; 778-799 AND 1325-1345.
CC RP SPANNE-SE7 / ATCC 29145;
CC RC MEDLINE=93155143; Pubmed=8428988;
CC RX MEDLINE=93155143; Pubmed=8428988;
CC RA Pelanda R., Vancil M.A., Perego M., Plubelli L., Gallizzi A.,
CC Curti B., Zanetti G.;
CC RT "Isolation of a glutamate synthase genes of the diazotroph Azospirillum brasilense.
CC Cloning, sequencing, and analysis of functional domains.";
CC J Biol. Chem. 268:3099-3106(1993).
CC [2]
CC RN SEQUENCE OF 834-927 FROM N.A.
CC RP STRAIN=RG;
CC RC MEDLINE=94075244; Pubmed=7902833;
CC RX MEDLINE=94075244; Pubmed=7902833;
CC RA Mandal A.K., Ghosh S.;
CC RT "Structural studies on the subunits of glutamate synthase from
CC Azospirillum brasilense.";
CC RL Biochim. Biophys. Acta 1039:374-377(1990).
CC CC -1- CATALYTIC ACTIVITY: 2 L-glutamate + NADP(+) = L-glutamine + 2-
CC oxoglutarate + NADPH.
CC CC -1- COFACTOR: BINDS A 3FE-4S CLUSTER; FAD AND FMN.
CC CC -1- PATHWAY: NITROGEN METABOLISM, GLUTAMATE BIOSYNTHESIS.
CC CC THE CATALYZED REACTION BRINGS TOGETHER THE NITROGEN AND
CC CARBON METABOLISM.
CC CC -1- SUBUNIT: AGGREGATE OF 4 CATALYTICAL ACTIVE HETERODIMERS,
CC CONSISTING OF A LARGE AND A SMALL SUBUNIT.
CC CC -1- MISCELLANEOUS: GLUTAMINE BINDS TO THE LARGE SUBUNIT AND TRANSFERS
CC THE AMIDO GROUP TO 2-OXO-GLUTAMATE THAT APPARENTLY BINDS TO THE
CC SMALL SUBUNIT.
CC CC -1- SIMILARITY: TO OTHER GLUTAMATE SYNTHASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF192408; AAA22179.1; -.
CC DR EMBL: X71632; CAA50639.1; -.
CC PIR: B46602; B46602.
CC DR InterPro: IPR002489; DUF14.
CC DR InterPro: IPR003009; FMN-enzyme.
CC DR InterPro: IPR002932; Glu_synthase.
CC DR Pfam: PF01493; DUF14; 1.
CC DR Pfam: PF01645; Glu_synthase; 1.
CC KW Oxidoreductase; Iron-sulfur; 3Fe-4S; Flavoprotein; FAD; FMN; NADP;
CC Glutamate biosynthesis; Zymogen.
CC FT PROPEP 1 36
CC FT CHAIN 37 1515
CC FT NP_BIND 1085 1142 FMN (BY SIMILARITY).
CC FT METAL 1138 1138 IRON-SULFUR (3FE-4S) (BY SIMILARITY).
CC FT METAL 1144 1144 IRON-SULFUR (3FE-4S) (BY SIMILARITY).
CC FT METAL 1149 1149 IRON-SULFUR (3FE-4S) (BY SIMILARITY).
CC FT NP_BIND 1086 1142 FMN (BY SIMILARITY).
CC SEQENCE 1515 AA; 166018 MW; ED91321BDFCE92 CRC64;

Query Match 58.3%; Score 35; DB 1; Length 1515;
Best Local Similarity 70.0%; Pred. NO. 78;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 AEGLDTORFS 11
   11111111

```

Tue Apr 15 10:21:36 2003

us-09-521-527c-1.open.rsp

Page 10

Db 683 AEGDTHFA 692

Search completed: April 15, 2003, 10:12:44
Job time : 13 secs

GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 15, 2003, 10:11:09 ; Search time 29 Seconds
(without alignments)
85.261 Million cell updates/sec

Title: us-09-521-527c-1

Perfect score: 60

Sequence: 1 AAEGDTRFSG 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: SP archaea:*
- 2: SP bacteria:*
- 3: SP fungi:*
- 4: SP human:*
- 5: SP invertebrate:*
- 6: SP mammal:*
- 7: SP mhc:*
- 8: SP organelle:*
- 9: SP phage:*
- 10: SP plant:*
- 11: SP rodent:*
- 12: SP virus:*
- 13: SP vertebrate:*
- 14: SP unclassified:*
- 15: SP_rvirus:*
- 16: SP_bacteriap:*
- 17: SP_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	198	4 Q13970	Q13970 homo sapien
2	60	100.0	235	4 Q960R6	Q960R6 homo sapien
3	60	100.0	235	4 Q8TAM8	Q8TAM8 homo sapien
4	48	80.0	235	6 Q9XSM6	Q9XSM6 salmifit sci
5	39	65.0	973	4 Q96D13	Q96D13 homo sapien
6	39	65.0	973	4 Q8TCG0	Q8TCG0 homo sapien
7	39	65.0	986	4 Q9P253	Q9P253 homo sapien
8	38	63.3	631	16 Q8X6X9	Q8X6X9 escherichia
9	38	63.3	817	2 Q8VUT2	Q8VUT2 pseudomonas
10	38	63.3	1208	12 Q9LIF8	Q9LIF8 white spot
11	38	63.3	1208	12 Q8VB56	Q8VB56 white spot
12	37	61.7	252	16 Q9PB53	Q9PB53 xylella fas
13	37	61.7	299	2 Q9RBF7	Q9RBF7 alcaligenes
14	37	61.7	363	16 Q8XXX5	Q8XXX5 raietonia s
15	37	61.7	370	10 Q9FKA8	Q9FKA8 arbidopsis
16	37	61.7	726	16 Q9PPH5	Q9PPH5 xylella fas

17	37	61.7	726	16 Q9PAJ1	Q9PAJ1 xylella fas
18	36	60.0	203	2 Q9X7L5	Q9X7L5 rhizobium m
19	36	60.0	218	10 Q9LD82	Q9LD82 oryza sativ
20	36	60.0	461	10 Q9FLK4	Q9FLK4 arbidopsis
21	36	60.0	988	16 Q9K4L9	Q9K4L9 streptomyces
22	35.5	59.2	292	10 Q8S0H6	Q8S0H6 oryza sativ
23	35	58.3	111	9 Q8SCF9	Q8SCF9 vibrio harv
24	35	58.3	219	16 Q9K4F1	Q9K4F1 streptomyces
25	35	58.3	233	11 Q9D4X6	Q9D4X6 mus musculu
26	35	58.3	26	16 Q8ZGM5	Q8ZGM5 yersinia pe
27	35	58.3	417	16 P96441	P96441 rhizobium m
28	35	58.3	505	16 Q9JMA0	Q9JMA0 neisseria m
29	35	58.3	514	17 Q8TVA9	Q8TVA9 methanopyru
30	35	58.3	523	17 Q974R9	Q974R9 sulfolobus
31	35	58.3	524	16 Q9A354	Q9A354 caulobacter
32	35	58.3	610	16 Q9CLU3	Q9CLU3 pasteurella
33	35	58.3	614	16 Q9PGP6	Q9PGP6 xylella fas
34	35	58.3	615	13 Q9D835	Q9D835 pediculus
35	35	58.3	631	16 Q8Z326	Q8Z326 salmonella
36	35	58.3	646	16 Q8ZL74	Q8ZL74 salmonella
37	35	58.3	646	16 Q8X871	Q8X871 salmonella
38	35	58.3	646	16 Q8XA87	Q8XA87 escherichia
39	35	58.3	664	16 Q8Z8C8	Q8Z8C8 yersinia pe
40	35	58.3	736	6 Q97490	Q97490 oryctolagus
41	35	58.3	1082	12 Q69514	Q69514 human herpe
42	35	58.3	1199	12 Q56301	Q56301 human herpe
43	35	58.3	1205	10 Q9FL35	Q9FL35 arbidopsis
44	35	58.3	1428	17 Q8TR62	Q8TR62 methanosarc
45	34.5	57.5	276	2 Q8RML5	Q8RML5 enterobacte

ALIGNMENTS

RESULT 1

ID	Q13970	PRELIMINARY;	PRT;	198 AA.
AC	Q13970;			
DT	01-NOV-1996 (TREMblrel. 01, Created)			
DT	01-NOV-1996 (TREMblrel. 01, Last sequence update)			
DT	01-DEC-2001 (TREMblrel. 19, Last annotation update)			
DE	CD8 antigen.			
GN	CD8A.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=89215302; PubMed=2496167;			
RA	Norment A.M., Lomborg N., Lacy E., Littman D.R.;			
RT	"Alternatively spliced mRNA encodes a secreted form of human CD8			
RT	alpha. Characterization of the human CD8 alpha gene.";			
KL	J. Immunol. 142:3312-3319(1989).			
DR	EMBL; M26315; AAA79218.1; -			
DR	EMBL; M26313; AAA79218.1; JOINED.			
DR	EMBL; M26314; AAA79218.1; JOINED.			
DR	HSSP; P01732; ICD8.			
DR	InterPro; IPR003596; IG_MHC.			
DR	InterPro; IPR003596; IG_V.			
DR	Pfam; PF00047; Ig; 1.			
DR	SMART; SM00406; IGV; 1.			
SO	SEQUENCE 198 AA; 21585 MW; BC89FA59F150598F CRC64;			

Query Match	100.0%;	Score 60;	DB 4;	Length 198;
Best Local Similarity	100.0%;	Pred. No. 0.001;		
Matches 12;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1 AAEGDTRFSG 12			
DB	80 AAEGDTRFSG 91			

RESULT 2

Q960R6 PRELIMINARY; PRT; 235 AA.
 AC Q960R6; 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Mutant CD8 alpha antigen.
 GN CD8A.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 RN NCBI_TaxID=9606;
 RP SEQUENCE FROM N.A.
 RX MEDLINE-21329138; PubMed-11435463;
 RA de La Calle-Martín O., Hernandez M., Ordí J., Casanltjana N.,
 RA Arostegui J.I., Caragol I., Ferrando M., Labrador M.,
 RA Rodriguez-Sanchez J.L., Espanol T.;
 RT "Familial CD8 deficiency due to a mutation in the CD8alpha gene."
 RL J. Clin. Invest. 108:117-117(2001).
 DR EMBL; AF039664; AAK72403.1; -.
 DR InterPro: IPR003006; 19_MHC.
 DR Pfam; PF00047; 19; 1.
 SQ SEQUENCE 235 AA; 25759 MW; FCCA347AAEF732BB CRC64;

Query Match 100.0%; Score 60; DB 4; Length 235;
 Best Local Similarity 100.0%; Pred. No. 0.0012;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAEGLDTRFSG 12
 DB 80 AAEGLDTRFSG 91

RESULT 3

Q87AW8 PRELIMINARY; PRT; 235 AA.
 AC Q87AW8; 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE CD8 antigen, alpha polypeptide (p32).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 RN NCBI_TaxID=9606;
 RP SEQUENCE FROM N.A.
 RC TISSUE-PANCREAS;
 RA Strausberg R.;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC025715; AAH25715.1; -.
 SQ SEQUENCE 235 AA; 25669 MW; FCDPC9B5472D26BB CRC64;

Query Match 100.0%; Score 60; DB 4; Length 235;
 Best Local Similarity 100.0%; Pred. No. 0.0012;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAEGLDTRFSG 12
 DB 80 AAEGLDTRFSG 91

RESULT 4

Q9XSM6 PRELIMINARY; PRT; 235 AA.
 AC Q9XSM6; 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE alpha chain precursor.
 GN Scirtureus (Common squirrel monkey).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Cebinae; Saimiri.
 RN NCBI_TaxID=9521;
 RP SEQUENCE FROM N.A.
 RC STRAIN=92039;
 RX MEDLINE-992939362; PubMed-10369934;
 RA Ureta-Vidal A., Garcia Z., Lemmonier F.A., Kazanj M.;
 RT "Molecular characterization of cDNAs encoding squirrel monkey (Saimiri
 scirtureus) CD8 alpha and beta chains."
 RL Immunogenetics 49:718-721(1999).
 DR EMBL; AJ130818; CAB41462.1; -.
 DR HSSP; P01732; 1CD8
 DR InterPro: IPR003599; 19.
 DR InterPro: IPR003006; 19_MHC.
 DR InterPro: IPR002965; P-rich_extensn.
 DR Pfam; PF00047; 19; 1.
 DR PRINTS; PR01217; PRICHEXTENS.
 DR SMART; SM00409; IG; 1.
 KW Signal.

FT SIGNAL 1 21 POTENTIAL.
 FT CHAIN 22 235 CD8 ALPHA CHAIN.
 SQ SEQUENCE 235 AA; 25728 MW; 055867CD503C268D CRC64;

Query Match 80.0%; Score 48; DB 6; Length 235;
 Best Local Similarity 81.8%; Pred. No. 0.28;

Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 AAEGLDTRFSG 12
 DB 81 ADGLDTRFSG 91

RESULT 5

Q96D13 PRELIMINARY; PRT; 580 AA.
 AC Q96D13; 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Hypothetical 65.0 kDa protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 RN NCBI_TaxID=9606;
 RP SEQUENCE FROM N.A.
 RC TISSUE-LUNG;
 RA Strausberg R.;
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC001513; AAH01513.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 580 AA; 64958 MW; 9FB02B4160595E32 CRC64;

Query Match 65.0%; Score 39; DB 4; Length 580;
 Best Local Similarity 66.7%; Pred. No. 45;

Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 AAEGLDTRFSG 12
 DB 235 AAEGLDTRFSG 246

RESULT 6

Q8TCG0 PRELIMINARY; PRT; 973 AA.
 AC Q8TCG0; 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Hypothetical 110.2 kDa protein.
 GN DKFZP762G147.
 OS Homo sapiens (Human).

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.
RC TISSUE=MELANOMA;
RA Oltjenweider B., Obermaier B., Mewes H.W., Well B., Wiemann S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL713725; CAD28515.1; -
KM Hypothetical protein.
SQ SEQUENCE 973 AA; 110185 MW; 0008E21D66E8CEFC CRC64;

Query Match
Best Local Similarity 65.0%; Score 39; DB 4; Length 973;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAEGLDTQRFSG 12
DB 235 AAECAEAGGFSG 246

RESULT 7
O9P253 PRELIMINARY; PRT; 986 AA.
AC O9P253; Q9H268;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE KIAA1475 protein (Vacuolar protein sorting protein 18)
DE (Fragment).
GN KIAA1475 OR VPS18.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.
RA Nagase T., Kikuno R., Ishikawa K., Hirosewa M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human
RT genes. XVII. The complete sequences of 100 new cDNA clones from brain
RT which code for large proteins in vitro.";
RL DNA Res. 7:143-150(2000).
RN [2]
RN SEQUENCE OF 14-986 FROM N.A.
RX MEDLINE=20277482; PubMed=10819331;
RA Huizing M., Didier A., Walenta J., Anikster Y., Gahl W.A., Kramer H.;
RT "Molecular cloning and characterization of human VPS18, VPS11, VPS16,
RT and VPS33.";
RL Gene 264:241-247(2001).
DR EMBL: AB040908; BAA9599.1; -
DR EMBL: AF308802; AAC34679.1; -
DR InterPro: IPR00547; Clathrin_repeat.
DR SMART: SM00299; CLH; 1.
FT NON_TER
SQ SEQUENCE 986 AA; 111475 MW; 4BA494BE6ED58C94 CRC64;

Query Match
Best Local Similarity 65.0%; Score 39; DB 4; Length 986;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAEGLDTQRFSG 12
DB 248 AAECAEAGGFSG 259

RESULT 8
O8X6X9 PRELIMINARY; PRT; 631 AA.
AC O8X6X9;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

```

```

DE Thiamin biosynthesis, pyrimidine moiety (Thiamin biosynthesis protein
DE Th1C).
GN TH1C OR Z5569 OR ECS4917.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G., III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick J.A.,
RA Posfal G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobbeck E.J., Davis N.W., Lam A., Dimlantia E.T., Polanousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsuda E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
DR EMBL: AE005631; AAG59191.1; -
DR EMBL: AP002567; BAB38340.1; -
DR InterPro: IPR002817; Th1C.
DR Pfam: PF01964; Th1C; 1.
DR ProDom: PD007048; Th1C; 1.
DR TIGRFAMs: TIGR00190; Th1C; 1.
KW Complete proteome.
SQ SEQUENCE 631 AA; 70814 MW; 5494CBBD9C036F2E CRC64;

Query Match
Best Local Similarity 63.3%; Score 38; DB 16; Length 631;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAEGLDTQRFSG 12
DB 123 ADDGDELRFSG 134

RESULT 9
O8VUT2 PRELIMINARY; PRT; 817 AA.
AC O8VUT2;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Beta-1,3-xylanase.
GN AXNB.
OS Pseudomonas sp. ND137.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=147640;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=ND137;
RA Aoki Y., Kitamura E., Myouga H., Kamei Y.;
RT "Protoplast production of red alga Porphyra yezoensis by recombinant
RT enzymes from Pseudomonas sp.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB063257; BAB9289.1; -
KM Xylan degradation; Hydrolase; Glycosidase.
SQ SEQUENCE 817 AA; 85650 MW; 6D453A7D6D202329 CRC64;

Query Match
Best Local Similarity 63.3%; Score 38; DB 2; Length 817;

```

RA Wang C.H., Lin J.Y., Kou G.H., Lo C.F.:
RT Identification and characterization of a shrimp white spot syndrome
RT virus (WSSV) gene that encodes a novel chimeric polypeptide of

RA de Souza A.P., Ten
RA Vallada H., Van SI
RA Zago M.A., Zatz M

```

RT      "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL      Nature 406:151-159(2000).
DR      EMBL; AE004041; AAF85090.1; -.
KW      Complete proteome.
SQ      SEQUENCE 252 AA; 28014 MW; 4EB9FCBE739F01FB CRC64;

Query Match
Best Local Similarity 61.7%; Score 37; DB 16; Length 252;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      1 AAEGLDTPRF 10
        ||| |||
DB      240 AAEGLDTPRY 249

RESULT 13
Q9RBF7
ID      Q9RBF7; PRELIMINARY; PRT; 299 AA.
AC      Q9RBF7;
DT      01-MAY-2000 (TrEMBLrel. 13, Created)
DT      01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT      01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE      Formate dehydrogenase.
GN      FDSR.
OS      Alcaligenes eutrophus (Ralstonia eutropha).
OC      Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC      Ralstonia.
OX      NCBI_TaxID=510;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=H16;
RX      MEDLINE=20032351; PubMed=10564479;
RT      Oh J.I., Bowlen B.;
RT      "Dual control by regulatory gene fdsR of the fds operon encoding the
RT      NAD(+)-linked formate dehydrogenase of Ralstonia eutropha.";
RL      Mol. Microbiol. 34:365-376(1999).
DR      EMBL; AA007718; CAB59983.1; -.
DR      InterPro: IPR003725; Mode.
DR      Pfam: PF02573; HTH_9; 1.
SQ      SEQUENCE 299 AA; 32319 MW; 43DD1AD3D40CA4BC CRC64;

Query Match
Best Local Similarity 61.7%; Score 37; DB 2; Length 299;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY      1 AAEGLDTPRFS 12
        || ||| |
DB      249 AARGIDTARIEG 260

RESULT 14
O8XXK5
ID      O8XXK5; PRELIMINARY; PRT; 363 AA.
AC      O8XXK5;
DT      01-MAR-2002 (TrEMBLrel. 20, Created)
DT      01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT      01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE      Probable 3-isopropylmalate dehydrogenase oxidoreductase protein
DE      (EC 1.1.1.85).
GN      LEU1 OR RSC1988 OR RS03417.
OS      Ralstonia solanacearum (Pseudomonas solanacearum).
OC      Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC      Ralstonia.
OX      NCBI_TaxID=305;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=GM11000;
RX      MEDLINE=21681879; PubMed=11823852;
RA      Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA      Arlat M., Billault A., Brotlier P., Camus J.C., Catolico L.,
RA      Chandler M., Choise N., Claudel-Renard C., Cunnac S., Demange N.,
RA      Gaspin C., Lavie M., Molzan A., Robert C., Saurin W., Schlex T.,
RA      Stigter P., Thebaud P., Whalen M., Wincker P., Levy M.,

```

```

RA      Weissenbach J., Boucher C.A.;
RT      "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL      Nature 415:497-502(2002).
DR      EMBL; AL646067; CAD15690.1; -.
DR      InterPro: IPR001804; Isoch.
DR      InterPro: IPR004429; LeuB.
DR      Pfam: PF00180; Isoch; 1.
DR      TIGRfams: TIGR00169; LeuB; 1.
DR      PROSITE: PS00470; IDH_IMDH; 1.
KW      Oxidoreductase; Complete proteome.
SQ      SEQUENCE 363 AA; 39128 MW; 6343123C09817ACE CRC64;

Query Match
Best Local Similarity 61.7%; Score 37; DB 16; Length 363;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      1 AAEGLDTPRFS 11
        || ||| ||
DB      161 AREGFTWRY 171

RESULT 15
Q9FKA8
ID      Q9FKA8; PRELIMINARY; PRT; 370 AA.
AC      Q9FKA8;
DT      01-MAR-2001 (TrEMBLrel. 16, Created)
DT      01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT      01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE      Retroelement pol polyprotein-like.
OS      Arabidopsis thaliana (Mouse-ear cress).
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC      eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX      NCBI_TaxID=3702;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=COLUMBIA;
RX      MEDLINE=98403884; PubMed=9734815;
RA      Kotani H., Nakamura Y., Sato S., Asamizu E., Kaneko T., Miyajima N.,
RA      Tabata S.;
RT      "Structural analysis of Arabidopsis thaliana chromosome 5. VI.
RT      Sequence features of the regions of 1,367,185 bp covered by 19
RT      physically assigned P1 and TAC clones.";
RL      DNA Res. 5:203-216(1998).
DR      EMBL; AB012243; BAB08885.1; -.
DR      InterPro: IPR005162; Retrotrans.-gag.
DR      InterPro: IPR001878; Znf_CCHC.
DR      Pfam: PF03732; Retrotrans.-gag; 1.
DR      Pfam: PF00098; Zf-CCHC; 1.
DR      SMART; SM00343; Znf_C2HC; 1.
KW      Polyprotein.
SQ      SEQUENCE 370 AA; 40618 MW; 3EBD9764902A6C17 CRC64;

Query Match
Best Local Similarity 61.7%; Score 37; DB 10; Length 370;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      4 GLDTPRFS 12
        ||| ||| ||
DB      201 GLDSARFS 209

```

Search completed: April 15, 2003, 10:13:21
 Job time : 32 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 15, 2003, 10:11:04 ; Search time 72 Seconds
(without alignments)
22.208 Million cell updates/sec

Title: us-09-521-527c-1
Perfect score: 60
Sequence: 1 AAEGLDTRFSG 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq.101002.*
1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	60	100.0	12	21	Human CD8 antigen
2	60	100.0	20	21	Human CD8 antigen
3	60	100.0	114	17	CD8 alpha chain
4	60	100.0	114	21	Extracellular domain
5	60	100.0	121	20	Expressed fragment
6	60	100.0	235	15	Sequence of human
7	60	100.0	235	20	Human CD8 alpha pr
8	60	100.0	273	15	Human single chain
9	60	100.0	376	21	Secreted modified
10	54	90.0	11	21	Human CD8 antigen

11	40	66.7	8	17	AA98502	CD8 alpha chain re
12	40	66.7	8	21	AA943646	Fragment of a huma
13	40	66.7	10	17	AA98509	CD8 antigenist #4.
14	40	66.7	10	21	AA943649	Amino acid sequenc
15	39	65.0	80	22	AA423815	Human EST encoded
16	39	65.0	973	23	AB97191	Novel human protei
17	39	65.0	973	23	AB905701	Human intracellular
18	38	63.3	374	21	AA971530	Corn POP2/CAPI tra
19	38	63.3	1204	22	AA685003	Shrimp white spot
20	36	60.0	87	21	AA964914	Human 5' EST relat
21	36	60.0	153	21	AA964914	Human 5' EST relat
22	36	60.0	373	21	AA966668	Arabidopsis thalia
23	36	60.0	420	21	AA966667	Arabidopsis thalia
24	36	60.0	461	21	AA966666	Arabidopsis thalia
25	36	60.0	553	18	AA935856	Human CD45 for use
26	35	58.3	278	21	AA932698	Eucaalyptus grandis
27	35	58.3	504	23	AA017865	Pyrin domain conta
28	35	58.3	505	21	AA975701	Neisseria meningit
29	35	58.3	505	23	AA972960	Neisseria meningit
30	35	58.3	505	23	AA972969	Neisseria meningit
31	35	58.3	505	23	AA972982	Neisseria meningit
32	35	58.3	589	19	AA960667	E.coli cold shock
33	35	58.3	624	23	AA914675	Corynebacterium q1
34	35	58.3	646	22	AA934718	E. coli cellular p
35	35	58.3	646	22	AA938176	Salmonella typhi c
36	35	58.3	655	23	AB907264	Human ATRG polylep
37	35	58.3	732	22	AA981024	C glutamicum prote
38	35	58.3	736	22	AA962880	Membrane bound tra
39	35	58.3	736	23	AA978362	Cell differentiat
40	35	58.3	1006	22	AB924871	Novel human diagno
41	35	58.3	1038	22	AB924868	Novel human diagno
42	34	56.7	49	21	AB932691	Eucaalyptus grandis
43	34	56.7	68	23	AB908088	Human ORFX protein
44	34	56.7	105	21	AA971535	Wheat POP2/CAPI tr
45	34	56.7	301	22	AB917511	Novel human diagno

ALIGNMENTS

RESULT 1
ID AAB24101 standard; peptide: 12 AA.
XX AAB24101;
AC AAB24101;
XX 23-JAN-2001 (first entry)
DT 23-JAN-2001 (first entry)
XX
XX Human CD8 antigenic peptide.
DE Human CD8 antigenic peptide.
XX
XX Human; CD8; antibody; antigenic; viral infection; tumour;
KW CD8+ cell; peripheral mononuclear blood cell.
XX
XX Homo sapiens.
OS
XX
XX WO200053305-A1.
PN
XX
XX 21-SEP-2000.
PD
XX
XX 08-MAR-2000; 2000WO-US05898.
PE
XX
XX 12-MAR-1999; 99US-0124253.
PR
XX
XX (ORTH) ORTHO-MCNEL PHARM INC.
PA
XX
XX Leturcq D;
PI
XX
XX WPI; 2000-618917/59.
DR
XX
XX Isolating CD8+ cells from peripheral mononuclear blood cells for use as
PT vehicles for combating viral infections and tumors, comprises using a
PT CD8 specific antibody that does not activate the cells -
XX

PS Claim 1; Page 15; 21pp; English.

XX The present invention describes a method for isolating CD8+ cells (1)

CC from peripheral mononuclear blood cells (PBMC) using an antibody which

CC specifically binds to the sequence AAEGLDTQRFSG (1) or a portion of (1)

CC on CD8 molecules present on the surface of CD8+ cells but does not

CC activate the CD8+ cells once bound. Also described are: (1) a hybridoma

CC cell line (II) which produces a monoclonal antibody that specifically

CC binds to CD8 molecules present on the surface of CD8+ cells but does not

CC activate the CD8+ cells; (2) a monoclonal antibody (III) produced by

CC (II); (3) a polypeptide comprising (1), useful for generating (III);

CC (4) a population of CD8+ cells isolated by (1); and (5) a kit, comprising

CC (III) and an agent which causes the dissociation of CD8+ cell-antibody

CC complex; (1) is useful for isolating human CD8+ cells from a sample of

CC isolated peripheral mononuclear blood cells. The isolated CD8+ cells have

CC importance as vehicles for combating viral infections and tumours.

XX

SQ Sequence 12 AA:

Query Match 100.0%; Score 60; DB 21; Length 12;

Best Local Similarity 100.0%; Pred. No. 1.4e-05;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAEGLDTQRFSG 12

DB 1 AAEGLDTQRFSG 12

RESULT 2

ID AAB24103 standard; peptide; 20 AA.

AC AAB24103;

DT 23-JAN-2001 (first entry)

DE Human CD8 antigenic peptide #1.

XX Human; CD8; antibody; antigenic; viral infection; tumour;

KW CD8+ cells; peripheral mononuclear blood cell.

OS Homo sapiens.

XX WO200055305-A1.

PN 21-SEP-2000.

PD 08-MAR-2000; 2000WO-US05898.

PF 12-MAR-1999; 98US-0124253.

XX (ORTH) ORTHO-MCNEIL PHARM INC.

PA Leturcq D;

PI WPI: 2000-618917/59.

DR Isolating CD8+ cells from peripheral mononuclear blood cells for use as

PT vehicles for combating viral infections and tumors, comprises using a

PT CD8 specific antibody that does not activate the cells

XX

PS Disclosure: Page 8; 21pp; English.

XX The present invention describes a method for isolating CD8+ cells (1)

CC from peripheral mononuclear blood cells (PBMC) using an antibody which

CC specifically binds to the sequence AAEGLDTQRFSG (1) or a portion of (1)

CC on CD8 molecules present on the surface of CD8+ cells but does not

CC activate the CD8+ cells once bound. Also described are: (1) a hybridoma

CC cell line (II) which produces a monoclonal antibody that specifically

CC binds to CD8 molecules present on the surface of CD8+ cells but does not

CC activate the CD8+ cells; (2) a monoclonal antibody (III) produced by

CC (II); (3) a polypeptide comprising (1), useful for generating (III);

CC (4) a population of CD8+ cells isolated by (1); and (5) a kit, comprising

CC (III) and an agent which causes the dissociation of CD8+ cell-antibody

CC complex; (1) is useful for isolating human CD8+ cells from a sample of

CC isolated peripheral mononuclear blood cells. The isolated CD8+ cells have

CC importance as vehicles for combating viral infections and tumors. The

CC present sequence represents a human CD8 antigenic peptide, which is

CC given in the exemplification of the present invention.

XX

SQ Sequence 20 AA:

Query Match 100.0%; Score 60; DB 21; Length 20;

Best Local Similarity 100.0%; Pred. No. 2.5e-05;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAEGLDTQRFSG 12

DB 5 AAEGLDTQRFSG 16

RESULT 3

ID AAR98514 standard; peptide; 114 AA.

AC AAR98514;

DT 04-MAR-1997 (first entry)

DE CD8 alpha chain.

XX CD8 antagonist; inhibitor; T-cell activation; human; CD8 alpha chain;

KW thymic differentiation; transplantation; bone marrow; liver; lung;

KW kidney; cornea; skin graft; graft versus host disease; therapy.

XX Homo sapiens.

XX MO9622106-A1.

PN 25-JUL-1996.

PD 17-JAN-1996; 96WO-US00310.

PF 17-JAN-1995; 95US-0372952.

XX (UYJE-) UNIV JEFFERSON THOMAS.

PA Choksi S, Huang Z, Jameson BA, Korggold R;

PI WPI: 1996-354307/35.

DR CD8 antagonist peptide(s) - used for inhibiting T cell activation,

PT partic. for treating transplant rejection or graft versus host

PT disease

XX

PS Disclosure: Page 27-28; 44pp; English.

XX This sequence represents the human CD8 alpha chain. CD8 plays a major

CC role in the activation of mature T-cells, and in the thymic

CC differentiation process that leads to CD8 expression. CD8 is expressed

CC either as a homodimer (containing two alpha chains) or as a heterodimer

CC (an alpha and a beta chain). The CD8 CD82-like region is involved in

CC regulating T-cell activation. The antagonists of the invention contain

CC a fragment of this sequence (see AAR98501-R98505), and comprise a

CC molecular surface similar to at least a portion of human CD8 molecular

CC surface around the site of one of these peptides. The compounds of the

CC invention compete with CD8 so as to inhibit T-cell activation. They can

CC also be used to treat an individual who is about to undergo, is

CC undergoing, or has undergone a transplantation procedure such as bone

CC marrow, liver, heart, kidney, lung, islets, or cornea transplantation, or

CC skin grafts. The compounds can also be used to treat an individual

CC suspected of suffering from, or susceptible to graft versus host

CC disease.

XX

SQ Sequence 114 AA:

Query Match 100.0%; Score 60; DB 17; Length 114;
 Best Local Similarity 100.0%; Pred. No. 0.00017;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAEGLDTRFSG 12
 |||||
 Db 59 AAEGLDTRFSG 70

RESULT 4
 ID AAY43644
 AC AAY43644 standard; peptide; 114 AA.
 XX
 AC AAY43644;
 XX
 DT 11-FEB-2000 (first entry)
 XX
 DE Extracellular domain of the alpha chain of human CD8.
 XX
 KM Extracellular domain; alpha chain; human CD8; T cell activation;
 KW cytotoxic T lymphocyte activation; CTL activation;
 KM graft versus host disease; CD8 antagonist; transplant;
 KM allogeneic bone marrow transplant.
 XX
 OS Homo sapiens.
 XX
 PN WO954345-A1.
 XX
 PD 28-OCT-1999.
 XX
 PF 21-APR-1999; 99WO-US08814.
 XX
 PR 21-APR-1998; 98US-0082436.
 XX
 PA (UYJE-) UNIV JEFFERSON THOMAS.
 XX
 PI Korgold R, Huang Z, Choksi S;
 XX
 DR WPI: 2000-013224/01.
 XX
 PT Novel CD8 antagonists used to inhibit cytotoxic T lymphocytes and
 PT prevent immune responses -
 XX
 PS Disclosure: Page 47-48; 67pp; English.
 XX
 CC The present sequence represents the extracellular domain of the alpha
 CC chain of human CD8. Regions of the alpha chain have been identified
 CC which interact with other molecules in the mediation of immune responses
 CC and activation of T cells. The specification describes peptides
 CC that are CD8 antagonists, and that compete with these CD8 regions in
 CC intermolecular interactions that involve CD8 which are associated with
 CC cytotoxic T lymphocyte (CTL) activation. These peptides comprise a
 CC molecular surface that corresponds to a molecular surface of human CD8
 CC at amino acids 73-76, 38-46, 53-56, 60-67 or 53-67. The peptides can be
 CC used in methods and compositions for therapeutically or prophylactically
 CC treating graft versus host disease in patients who have or about to
 CC undergo transplants, e.g. allogeneic bone marrow transplants. The
 CC methods and compositions are also used to treat individuals who are
 CC undergoing and/or who have undergone transplants, e.g. transplants of
 CC heart, liver, lung, islets or cornea, and skin grafts. The methods are
 CC also used to inhibit T cell activation in humans.
 CC
 XX
 SQ Sequence 114 AA;

Query Match 100.0%; Score 60; DB 21; Length 114;
 Best Local Similarity 100.0%; Pred. No. 0.00017;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAEGLDTRFSG 12
 |||||
 Db 59 AAEGLDTRFSG 70

RESULT 5
 ID AAY21689
 AC AAY21689 standard; Protein; 121 AA.
 XX
 AC AAY21689;
 XX
 DT 18-AUG-1999 (first entry)
 XX
 DE Expressed fragment of human CD8 alpha protein.
 XX
 KM T lymphocyte inhibition; CD8; immunosuppressive therapy; transplantation;
 KW autoimmune disease; allergy; asthma; viral infection; cytotoxic.
 KM corticosteroid; human; CD8 alpha protein.
 XX
 OS Homo sapiens.
 XX
 PN WO921576-A1.
 XX
 PD 06-MAY-1999.
 XX
 PF 28-OCT-1998; 98WO-GB03235.
 XX
 PR 28-OCT-1997; 97GB-0022779.
 XX
 PA (ISIS-) ISIS INNOVATION LTD.
 XX
 PI Gao GF, Gerth UC, Jakobsen BK, Sewell AK;
 XX
 DR WPI: 1999-385058/32.
 DR N-PSDB: AAX80968.
 XX
 PT Inhibiting activity of T cells against target cells useful for
 PT treating autoimmune diseases and allergy
 XX
 PS Claim 9; Fig 1B; 79pp; English.
 XX
 CC The invention describes a method for inhibiting activity of T lymphocytes
 CC against a target cell by treating the cell with a soluble form of a CD8
 CC molecule. The method is used as immunosuppressive therapy, e.g. in
 CC patients undergoing transplantation, but also for treating autoimmune
 CC diseases and allergy, e.g. exacerbation of asthma caused by viral
 CC infection. The CD8 protein inhibits cytotoxic T cells in vitro or in
 CC vivo, and its inhibitory activity can be adjusted by mutation. Treatment
 CC with CD8 protein provides more selective immunosuppression than use of
 CC corticosteroids. The present sequence represents the expressed fragment
 CC of human CD8 alpha protein.
 CC
 XX
 SQ Sequence 121 AA;

Query Match 100.0%; Score 60; DB 20; Length 121;
 Best Local Similarity 100.0%; Pred. No. 0.00018;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAEGLDTRFSG 12
 |||||
 Db 60 AAEGLDTRFSG 71

RESULT 6
 ID AAR9550
 AC AAR9550 standard; Protein; 235 AA.
 XX
 AC AAR9550;
 XX
 DT 20-AUG-1994 (first entry)
 XX
 DE Sequence of human CD8.
 XX
 KM gp39; T-cell antigen; CD40 ligand; B-cell proliferation;
 KM CD8; fusion protein; CD8; fusion protein.
 XX
 OS Homo sapiens.
 XX

PN EP585943-A.
 XX
 PD 09-MAR-1994.
 XX
 PF 03-SEP-1993; 93EP-0114153.
 XX
 PR 04-SEP-1992; 92US-0940605.
 XX
 PA (BRIM) BRISTOL-MYERS SQUIBB CO.
 XX
 PI Aruffo AA, Hollenbaugh D, Ledbetter JA, Aruffo A;
 DR WPI; 1994-076264/10.
 DR N-PSDB; AAQ57986.
 XX
 PT New nucleic acid encoding human gp39 T cell antigen - which is a
 PT ligand for the CD40 receptor, causing proliferation and
 PT differentiation of B cells and some cancer cells
 XX
 PS Disclosure; Fig 9; 39pp; English.
 XX
 CC The complete nucleic acid sequence of human gp39 (hgp39) protein
 CC (corresp. to CDNA) and the complete AA sequence of hgp39 are
 CC presented in AA057984 and AAR49548 respectively and contd. in plasmid
 CC CDM8-hgp39, deposited with the ATCC as E. coli, CDM8 MC1061/p3-hgp39
 CC and assigned accession No. 69050. The human T cell antigen gp39 is a
 CC ligand for the CD40 receptor. Soluble gp39 may be produced using the
 CC expression vector CD8-gp39. Chimeric genes may be constructed by
 CC fusing sequences encoding the extracellular domains of gp39 and
 CC CD8, pref. murine or human CD8 protein. Plasmid p3-shgp39 encoding
 CC the fusion protein of the extracellular domains from gp39 and CD8
 CC is deposited in E.coli as ATCC 69049.
 XX
 SQ Sequence 235 AA;

Query Match 100.0%; Score 60; DB 15; Length 235;
 Best Local Similarity 100.0%; Pred. No. 0.00039;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAEGLDTRFSG 12
 |||||||||
 DB 80 AAEGLDTRFSG 91

RESULT 7
 ID AAY21688
 XX AAY21688 standard; Protein; 235 AA.
 AC AAY21688;
 XX
 DT 18-AUG-1999 (first entry)
 XX
 DE Human CD8 alpha protein.
 XX
 KW T lymphocyte inhibition; CD8; immunosuppressive therapy; transplantation;
 KW autoimmune disease; allergy; asthma; viral infection; cytotoxic;
 KW corticosteroid; human; CD8 alpha protein.
 XX
 OS Homo sapiens.
 XX
 PN WO921576-A1.
 XX
 PD 06-MAY-1999.
 XX
 PE 28-OCT-1998; 98WO-GB03235.
 XX
 PR 28-OCT-1997; 97GB-0022779.
 XX
 PA (ISIS-) ISIS INNOVATION LTD.
 XX
 PI Gao GF, Gerth UC, Jakobsen BK, Sewell AK;
 XX WPI; 1999-385058/32.
 DR

DR N-PSDB: AAX80967.
 XX
 XX Inhibiting activity of T cells against target cells useful for
 PT treating autoimmune diseases and allergy
 XX
 PS Disclosure; Fig 1A; 79pp; English.

XX
 CC The invention describes a method for inhibiting activity of T lymphocytes
 CC against a target cell by treating the cell with a soluble form of a CD8
 CC molecule. The method is used as immunosuppressive therapy, e.g. in
 CC patients undergoing transplantation, but also for treating autoimmune
 CC diseases and allergy, e.g. exacerbation of asthma caused by viral
 CC infection. The CD8 protein inhibits cytotoxic T cells in vitro or in
 CC vivo, and its inhibitory activity can be adjusted by mutation. Treatment
 CC with CD8 protein provides more selective immunosuppression than use of
 CC corticosteroids. The present sequence represents a human CD8 alpha
 CC protein.
 XX
 SQ Sequence 235 AA;

Query Match 100.0%; Score 60; DB 20; Length 235;
 Best Local Similarity 100.0%; Pred. No. 0.00039;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAEGLDTRFSG 12
 |||||||||
 DB 80 AAEGLDTRFSG 91

RESULT 8
 ID AAR52862
 XX AAR52862 standard; Protein; 273 AA.
 AC AAR52862;
 XX
 DT 09-SEP-1994 (first entry)
 XX
 DE Human single chain CD8 in pPOW.

XX
 KW CD8; single chain antibody; target binding polypeptide;
 KW antibody engineering; humanized antibody; pPOW; vector.
 XX
 OS Homo sapiens.
 XX
 PN WO9407921-A.
 XX
 PD 14-APR-1994.
 XX
 PE 24-SEP-1993; 93WO-AU00491.
 XX
 PR 25-SEP-1992; 92AU-0004973.
 XX
 PA (CSIR) COMMONWEALTH SCI & IND RES ORG.

PI Atwell JL, Colman PM, Hudson PJ, Irving RA, Kortt A;
 PI Leib M, Malytil, Power BE;
 DR N-PSDB; AAQ62954.
 XX
 DR WPI; 1994-135515/16.
 XX
 PT New target-binding polypeptide(s) used for diagnosis, etc. -
 PT having a stable core polypeptide region with at least one
 PT target-binding region covalently attached, opt. mutated to alter
 PT specificity, etc.
 XX
 PS Disclosure; Page 38; 67pp; English.

XX Human single-chain CD8 construct, designed for expression in a
 CC bacterial secretion vector such as pPOW, has the DNA sequence
 CC given in AAQ62954, encoding the protein given in AAR52862.
 XX
 SQ Sequence 273 AA;

RESULT 11
 AAR98502
 ID AAR98502 standard; peptide: 8 AA.
 XX
 AC AAR98502;
 XX
 DT 04-MAR-1997 (first entry)
 XX
 DE CD8 alpha chain residues 60-67.
 XX
 KW CD8 antagonist; inhibitor; T-cell activation; human; CD8 alpha chain;
 KW thymic differentiation; transplantation; bone marrow; liver; heart; lung;
 KW kidney; cornea; skin graft; graft versus host disease; therapy.
 XX
 OS Homo sapiens.
 XX
 PN WO962106-A1.
 XX
 PD 25-JUL-1996.
 XX
 PF 17-JAN-1996; 96WO-US00310.
 XX
 PR 17-JAN-1995; 95US-0372952.
 XX
 PA (UYJE-) UNIV JEFFERSON THOMAS.
 XX
 PI Choksi S, Huang Z, Jameson BA, Korgold R;
 DR WPI: 1996-354307/35.
 XX
 PT CD8 antagonist peptide(s) - used for inhibiting T cell activation,
 PT partic. for treating transplant rejection or graft versus host
 PT disease
 XX
 PS Claim 2; Page 28; 44pp; English.
 XX
 CC AAR98501-R98505 represent components of a CD8 antagonist of the
 CC invention. These sequences are fragments of the human CD8 alpha chain
 CC molecule (see AAR98514). This sequence corresponds to residues 60-67 of
 CC the human CD8 sequence. CD8 plays a major role in the activation of
 CC mature T-cells, and in the thymic differentiation process that leads to
 CC CD8 expression. CD8 is expressed either as a homodimer (containing two
 CC alpha chains) or as a heterodimer (an alpha and a beta chain). The CD8
 CC CDR2-like region is involved in regulating T-cell activation. The
 CC antagonists of the invention comprise a molecular surface similar to at
 CC least a portion of human CD8 molecular surface around the site of one of
 CC these peptides. The compounds of the invention compete with CD8 so as to
 CC inhibit T-cell activation. They can also be used to treat an individual
 CC who is about to undergo, is undergoing, or has undergone a
 CC transplantation procedure such as bone marrow, liver, heart, kidney,
 CC lung, islets, or cornea transplantation, or skin grafts. The compounds
 CC can also be used to treat an individual suspected of suffering from, or
 CC susceptible to graft versus host disease.
 XX
 SQ Sequence 8 AA;
 XX
 QY Query Match 66.7%; Score 40; DB 17; Length 8;
 DB Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 2 AEGIDTOR 9
 DB 1 AEGIDTOR 8
 XX
 AC AAY43646
 ID AAY43646 standard; peptide: 8 AA.
 XX
 AC AAY43646;
 XX

DT 11-FEB-2000 (first entry)
 XX
 DE Fragment of a human CD8 antagonist peptide.
 XX
 KW Extracellular domain; alpha chain; human CD8; T cell activation;
 KW cytotoxic T lymphocyte activation; CTL activation;
 KW graft versus host disease; CD8 antagonist; transplant;
 KW allogeneic bone marrow transplant.
 XX
 OS Homo sapiens.
 XX
 PN WO954345-A1.
 XX
 PD 28-OCT-1999.
 XX
 PF 21-APR-1999; 99WO-US008814.
 XX
 PR 21-APR-1998; 98US-0082436.
 XX
 PA (UYJE-) UNIV JEFFERSON THOMAS.
 XX
 PI Korgold R, Huang Z, Choksi S;
 DR WPI: 2000-013224/01.
 XX
 PT Novel CD8 antagonists used to inhibit cytotoxic T lymphocytes and
 PT prevent immune responses
 XX
 PS Disclosure; Page 48; 67pp; English.
 XX
 CC The present sequence represents a fragment of a human CD8 antagonist
 CC peptide. The specification describes peptides that are CD8 antagonists,
 CC and that compete with regions within the extracellular domain of the
 CC alpha chain of CD8 regions in intermolecular interactions that involve
 CC CD8 which are associated with cytotoxic T lymphocyte (CTL) activation.
 CC These peptides comprise a molecular surface that corresponds to a
 CC molecular surface of human CD8 at amino acids 73-76, 38-46, 53-56,
 CC 60-67 or 53-67. The peptides can be used in methods and compositions
 CC for therapeutically or prophylactically treating graft versus host
 CC disease in patients who have or about to undergo transplants, e.g.
 CC allogeneic bone marrow transplants. The methods and compositions are
 CC also used to treat individuals who are undergoing and/or who have
 CC undergone transplants, e.g. transplants of heart, lung, islets
 CC or cornea, and skin grafts. The methods are also used to inhibit T cell
 CC activation in humans.
 XX
 SQ Sequence 8 AA;
 XX
 QY Query Match 66.7%; Score 40; DB 21; Length 8;
 DB Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 2 AEGIDTOR 9
 DB 1 AEGIDTOR 8
 XX
 AC AAR98509
 ID AAR98509 standard; peptide: 10 AA.
 XX
 AC AAR98509;
 XX
 DT 04-MAR-1997 (first entry)
 XX
 DE CD8 antagonist #4.
 XX
 KW CD8 antagonist; inhibitor; T-cell activation; human; CD8 alpha chain;
 KW thymic differentiation; transplantation; bone marrow; liver; heart; lung;
 KW kidney; cornea; skin graft; graft versus host disease; therapy.
 XX
 OS Synthetic.
 XX

PN WO9622106-A1.
 PD 25-JUL-1996.
 XX
 PF 17-JAN-1996; 96WO-US00310.
 XX
 PR 17-JAN-1995; 95US-0372952.
 XX
 PA (UYJE-) UNIV JEFFERSON THOMAS.
 XX
 PI Choksi S, Huang Z, Jameson BA, Korngold R;
 XX WPI; 1996-354307/35.
 DR
 XX
 PT CD8 antagonist peptide(s) - used for inhibiting T cell activation,
 PT partic. for treating transplant rejection or graft versus host
 PT disease
 XX
 PS Claim 16; Page 29; 44pp; English.
 XX
 CC AAR98506-R98511 represent CD8 antagonists of the invention. CD8 plays a
 CC major role in the activation of mature T-cells, and in the thymic
 CC differentiation process that leads to CD8 expression. CD8 is expressed
 CC either as a homodimer (containing two alpha chains) or as a heterodimer
 CC (an alpha and a beta chain). The CD8 CDR2-like region is involved in
 CC regulating T-cell activation. The antagonists of the invention comprise
 CC a molecular surface similar to at least a portion of human CD8 molecular
 CC surface around the site of one of these peptides. The compounds of the
 CC invention compete with CD8 so as to inhibit T-cell activation. They can
 CC also be used to treat an individual who is about to undergo, is
 CC undergoing, or has undergone a transplantation procedure such as bone
 CC marrow, liver, heart, kidney, lung, islets, or cornea transplantation,
 CC or skin grafts. The compounds can also be used to treat an individual
 CC suspected of suffering from, or susceptible to graft versus host
 CC disease.
 CC
 SQ Sequence 10 AA;
 Query Match 66.7%; Score 40; DB 17; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.15;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 AEGLDTOR 9
 ID | | | | | | | |
 DB 2 AEGLDTOR 9
 DE
 AC AAY43649;
 XX
 XX
 DT 11-FEB-2000 (first entry)
 XX
 DE Amino acid sequence of a human CD8 antagonist peptide.
 XX
 XX Extracellular domain; alpha chain; human CD8; T cell activation;
 KW cytotoxic T lymphocyte activation; CTL activation;
 KW graft versus host disease; CD8 antagonist; transplant;
 KW allogeneic bone marrow transplant; cyclic.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 XX
 PF Key Location/Qualifiers
 FT Disulfide-bond 1..10
 XX
 XX WO9554345-A1.
 XX
 PD 28-OCT-1999.
 XX
 PF 21-APR-1999; 99WO-US08814.

XX
 PR 21-APR-1998; 98US-0082436.
 XX
 PA (UYJE-) UNIV JEFFERSON THOMAS.
 XX
 PI Korngold R, Huang Z, Choksi S;
 XX WPI; 2000-013224/01.
 DR
 XX
 PT Novel CD8 antagonists used to inhibit cytotoxic T lymphocytes and
 PT prevent immune responses -
 XX
 PS Disclosure; Page 49; 67pp; English.
 XX
 CC The present sequence represents a human CD8 antagonist peptide of the
 CC invention. The specification describes peptides that are CD8 antagonists,
 CC and that compete with regions within the extracellular domain of the
 CC alpha chain of CD8 regions in intermolecular interactions that involve
 CC CD8 which are associated with cytotoxic T lymphocyte (CTL) activation.
 CC These peptides comprise a molecular surface that corresponds to a
 CC molecular surface of human CD8 at amino acids 73-76, 38-46, 53-56,
 CC 60-67 or 53-67. The peptides can be used in methods and compositions
 CC for therapeutically or prophylactically treating graft versus host
 CC disease in patients who have or about to undergo transplants, e.g.
 CC allogeneic bone marrow transplants. The methods and compositions are
 CC also used to treat individuals who are undergoing and/or who have
 CC undergone transplants, e.g. transplants of heart, liver, lung, islets
 CC or cornea, and skin grafts. The methods are also used to inhibit T cell
 CC activation in humans.
 CC
 SQ Sequence 10 AA;
 Query Match 66.7%; Score 40; DB 21; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.15;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 AEGLDTOR 9
 ID | | | | | | | |
 DB 2 AEGLDTOR 9
 DE
 AC AAM23815;
 XX
 XX
 DT 12-OCT-2001 (first entry)
 XX
 DE Human EST encoded protein SEQ ID NO: 1340.
 XX
 AC AAM23815 standard; Protein; 80 AA.
 XX
 XX
 DE Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
 KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
 KW diagnostics; forensic test; gene mapping; genetic disorder;
 KW biodiversity; gene therapy; nutrition.
 XX
 OS Homo sapiens.
 OS
 PN WO200154477-A2.
 XX
 PD 02-AUG-2001.
 XX
 PF 25-JAN-2001; 2001WO-US02687.
 XX
 XX
 PR 25-JAN-2000; 2000US-0491404.
 PR 17-JUL-2000; 2000US-0617746.
 PR 03-AUG-2000; 2000US-0631451.
 PR 15-SEP-2000; 2000US-0663870.
 XX
 XX (HYSE-) HYSEQ INC.
 PA
 XX Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
 PI Cao Y, Drmanac RA, Zhang J, Werhman T;

XX WPI: 2001-476164/51.
DR N-PSDB: AAH98474.
XX

XX Isolated polypeptide for treatment of diseases, diagnostics, raising
PT antibodies and research use -
XX

PS Claim 20; Page 948; 1275pp; English.

XX The present invention provides the protein and coding sequences of novel
CC proteins from a variety of organisms, including human, dog, cat, horse,
CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
CC urchin and tomato. These were derived from expressed sequence tags (ESTs)
CC from the organism of interest. They can be used in diagnostics,
CC forensics, gene mapping, identification of mutations, to assess
CC biodiversity and for nutritional purposes. The present sequence is a
CC protein of the invention.
XX

XX SQ Sequence 80 AA:

Query Match

Best Local Similarity 65.0%; Score 39; DB 22; Length 80;

Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 AAEGIDTORFS 11

I ||:|||||

DB 4 ACTGINTORFS 14

Search completed: April 15, 2003, 10:12:26
Job time : 73 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 15, 2003, 10:13:29 ; Search time 38 Seconds
(without alignments)
19.306 Million cell updates/sec

Title: US-09-521-527c-1
Perfect score: 60
Sequence: 1 AAEGLDTQRFSG 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 248812 seqs, 61136040 residues

Total number of hits satisfying chosen parameters: 248812

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- 1: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep.*
 - 2: /cgn2_6/ptodata/2/pubppaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/2/pubppaa/PCTUS_PUBCOMB.pep.*
 - 8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep.*
 - 10: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep.*
 - 12: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep.*
 - 13: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep.*
 - 14: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35	58.3	624	10	US-09-963-790A-2
2	35	58.3	646	10	US-09-815-242-10311
3	35	58.3	646	10	US-09-815-242-13769
4	35	58.3	655	9	US-10-066-521-13
5	35	58.3	732	9	US-09-738-626-4778
6	34	56.7	87	10	US-09-955-502-7
7	34	56.7	87	10	US-09-955-502-8
8	34	56.7	91	10	US-09-955-502-5
9	33	55.0	328	9	US-10-260-877-66
10	33	55.0	372	10	US-09-741-669-367
11	33	55.0	372	10	US-09-912-020-375
12	33	55.0	375	10	US-09-925-300-1674
13	33	55.0	550	10	US-09-895-828-452
14	33	55.0	635	10	US-09-815-242-13814
15	33	55.0	689	9	US-09-738-626-3675
16	33	55.0	2789	10	US-09-801-574-57
17	32	53.3	222	9	US-09-738-626-5078
18	32	53.3	241	9	US-10-102-806-610
19	32	53.3	260	9	US-09-738-626-4051

20	32	53.3	318	9	US-09-738-626-3557	Sequence 3557, Ap
21	32	53.3	467	9	US-09-738-626-6965	Sequence 6965, Ap
22	32	53.3	613	10	US-09-815-242-10990	Sequence 10990, A
23	32	53.3	806	9	US-09-815-242-11057	Sequence 11057, A
24	31.5	52.5	469	9	US-09-815-242-11057	Sequence 11057, A
25	31	51.7	22	10	US-09-896-888A-35	Sequence 35, Appl
26	31	51.7	23	10	US-09-896-888A-33	Sequence 33, Appl
27	31	51.7	27	10	US-09-896-888A-29	Sequence 29, Appl
28	31	51.7	28	10	US-09-896-888A-31	Sequence 31, Appl
29	31	51.7	43	10	US-09-896-888A-27	Sequence 27, Appl
30	31	51.7	87	10	US-09-955-502-6	Sequence 6, Appl1
31	31	51.7	88	10	US-09-955-502-9	Sequence 9, Appl1
32	31	51.7	90	10	US-09-955-502-10	Sequence 10, Appl1
33	31	51.7	116	9	US-09-738-626-6868	Sequence 6868, Ap
34	31	51.7	126	9	US-10-125-540-491	Sequence 491, Ap
35	31	51.7	126	10	US-09-764-870-491	Sequence 491, Ap
36	31	51.7	138	10	US-09-815-242-5416	Sequence 5416, Ap
37	31	51.7	142	10	US-09-815-242-12558	Sequence 12558, A
38	31	51.7	142	10	US-09-815-242-12561	Sequence 12561, A
39	31	51.7	204	9	US-09-738-626-4120	Sequence 4120, Ap
40	31	51.7	333	9	US-10-101-464A-561	Sequence 561, Ap
41	31	51.7	338	10	US-09-815-242-5727	Sequence 5727, Ap
42	31	51.7	341	10	US-09-820-044-2	Sequence 2, Appl1
43	31	51.7	341	10	US-09-815-242-12158	Sequence 12158, A
44	31	51.7	369	10	US-09-925-300-1070	Sequence 1070, Ap
45	31	51.7	370	10	US-09-862-027-20	Sequence 20, Appl

ALIGNMENTS

RESULT 1
US-09-963-790A-2
; Sequence 2, Application US/09963790A
; Patient No. US20020115161A1
; GENERAL INFORMATION:
; APPLICANT: FARMICK, Mike, et al.
; FILE OF INVENTION: NUCLEOTIDE SEQUENCES WHICH CODE FOR THE DEAD GENE
; TITLE REFERENCE: 032301 WD 230
; CURRENT APPLICATION NUMBER: US/09/963, 790A
; CURRENT FILING DATE: 2001-09-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 624
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-963-790A-2

Query Match 58.3%; Score 35; DB 10; Length 624;
Best Local Similarity 63.6%; Pred. No. 65;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAEGLDTQRFSG 11
DB 293 AAEGLDVERIS 303

RESULT 2
US-09-815-242-10311
; Sequence 10311, Application US/09815242
; Patient No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Traxick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes

```

FILE REFERENCE: ELTRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10311
LENGTH: 646
TYPE: PRT
ORGANISM: Escherichia coli
US-09-815-242-10311

```

```

Query Match      58.3%; Score 35; DB 10; Length 646;
Best Local Similarity 63.6%; Pred. No. 68;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

```

QY      1 AAEGLDQRF 11
      |||||:|
DB      322 AARGDVERIS 332

```

```

RESULT 3
US-09-815-242-13769
Sequence 13769, Application US/09815242
Patent No. US20020061369A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELTRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13769
LENGTH: 646
TYPE: PRT
ORGANISM: Salmonella typhi
US-09-815-242-13769

```

```

Query Match      58.3%; Score 35; DB 10; Length 646;
Best Local Similarity 63.6%; Pred. No. 68;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

```

QY      1 AAEGLDQRF 11
      |||||:|
DB      322 AARGDVERIS 332

```

```

RESULT 4
US-10-066-521-13
Sequence 13, Application US/10066521
Publication No. US20030027757A1
GENERAL INFORMATION:
APPLICANT: Bertin, John
APPLICANT: Wang, Weiye
APPLICANT: Blatcher, Maria
TITLE OF INVENTION: NOVEL MOLECULES OF THE PYRIN/NBS/LRR
FILE REFERENCE: 07334-334001
CURRENT APPLICATION NUMBER: US/10/066,521
CURRENT FILING DATE: 2002-06-25
PRIOR APPLICATION NUMBER: 60/318,645
PRIOR FILING DATE: 2001-09-10
PRIOR APPLICATION NUMBER: 60/265,231
PRIOR FILING DATE: 2001-01-31
NUMBER OF SEQ ID NOS: 25
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13
LENGTH: 655
TYPE: PRT
ORGANISM: Homo sapiens
US-10-066-521-13

```

```

Query Match      58.3%; Score 35; DB 9; Length 655;
Best Local Similarity 70.0%; Pred. No. 69;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

QY      1 AAEGLDQRF 10
      |||||:|
DB      417 AAEGLDQRF 426

```

```

RESULT 5
US-09-738-626-4778
Sequence 4778, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 4778
LENGTH: 732

```



```

; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4778
Query Match
Best Local Similarity 58.3%; Score 35; DB 9; Length 732;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 AEGIDTQRFSG 11
   |||||
DB 401 AARGLDVERIS 411

RESULT 6
US-09-955-502-7
; Sequence 7, Application US/09955502
; Patent No. US20020072118A1
; GENERAL INFORMATION:
; APPLICANT: Downs, Diana M.
; APPLICANT: Gralnick, Jeff A.
; TITLE OF INVENTION: Method for Preventing Superoxide Damage to Cells and
; FILE REFERENCE: 960296.97559
; CURRENT APPLICATION NUMBER: US/09/955,502
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: 60/234,588
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-955-502-7
Query Match
Best Local Similarity 56.7%; Score 34; DB 10; Length 87;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 AEGIDTQRFSG 12
   |||||
DB 14 AEGIDFQLYPG 24

RESULT 7
US-09-955-502-8
; Sequence 8, Application US/09955502
; Patent No. US20020072118A1
; GENERAL INFORMATION:
; APPLICANT: Downs, Diana M.
; APPLICANT: Gralnick, Jeff A.
; TITLE OF INVENTION: Method for Preventing Superoxide Damage to Cells and
; FILE REFERENCE: 960296.97559
; CURRENT APPLICATION NUMBER: US/09/955,502
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: 60/234,588
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Haemophilus ducreyi
US-09-955-502-8
Query Match
Best Local Similarity 56.7%; Score 34; DB 10; Length 87;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 AEGIDTQRFSG 12
   |||||
DB 14 AEGIDFQLYPG 24
```

```

RESULT 8
US-09-955-502-5
; Sequence 5, Application US/09955502
; Patent No. US20020072118A1
; GENERAL INFORMATION:
; APPLICANT: Downs, Diana M.
; APPLICANT: Gralnick, Jeff A.
; TITLE OF INVENTION: Method for Preventing Superoxide Damage to Cells and
; FILE REFERENCE: 960296.97559
; CURRENT APPLICATION NUMBER: US/09/955,502
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: 60/234,588
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Actinobacillus actinomycetemcomitans
US-09-955-502-5
Query Match
Best Local Similarity 56.7%; Score 34; DB 10; Length 91;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 AEGIDTQRFSG 12
   |||||
DB 14 AEGIDFQLYPG 24

RESULT 9
US-10-260-877-66
; Sequence 66, Application US/10260877
; Publication No. US20030021813A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Chovan, Linda E.
; APPLICANT: Hessler, Paul E.
; APPLICANT: Reich, Karl A.
; TITLE OF INVENTION: ESSENTIAL BACTERIA GENES AND GENOME
; TITLE OF INVENTION: SCANNING IN HAEMOPHILUS INFLUENZAE FOR THE IDENTIFICATION OF
; FILE REFERENCE: 6565.US.PI
; CURRENT APPLICATION NUMBER: US/10/260,877
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: US/09/649,145
; PRIOR FILING DATE: 2000-08-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 66
; LENGTH: 328
; TYPE: PRT
; ORGANISM: H. influenzae
US-10-260-877-66
Query Match
Best Local Similarity 55.0%; Score 33; DB 9; Length 328;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AEGIDTOR 9
   |||||
DB 286 AAGSDAQR 294

RESULT 10
US-09-741-669-367
; Sequence 367, Application US/09741669
; Patent No. US20020022718A1
; GENERAL INFORMATION:
; APPLICANT: Forsyth, R. Allyn
```

APPLICANT: Ohlsen, Karl L.
TITLE OF INVENTION: Genes identified as required for
FILE REFERENCE: ELITRA.009A
CURRENT APPLICATION NUMBER: US/09/741,669
CURRENT FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: US 60/173005
PRIOR FILING DATE: 1999-12-23
NUMBER OF SEQ ID NOS: 481
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 367
LENGTH: 372
TYPE: PRT
ORGANISM: Escherichia coli
US-09-741-669-367

Query Match 55.0%; Score 33; DB 10; Length 372;
Best Local Similarity 63.6%; Pred. No. 91;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 2 AEGIDTQRFSG 12
||| | | |
Db 250 AEGIDNVHFLG 260

RESULT 11
US-09-912-020-375
Sequence 375, Application US/09912020
Patent No. US20020045592A1
GENERAL INFORMATION:
APPLICANT: Zyskind, Judith
APPLICANT: Ohlsen, Karl L.
APPLICANT: Forsyth, R. Allan
APPLICANT: Froelich, Jamie M.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
TITLE OF INVENTION: GENES IDENTIFIED AS REQUIRED FOR PROLIFERATION IN
FILE REFERENCE: ELITRA.001DVI
CURRENT APPLICATION NUMBER: US/09/912,020
CURRENT FILING DATE: 2001-07-23
PRIOR APPLICATION NUMBER: 09/492,709
PRIOR FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: 60/117,405
PRIOR FILING DATE: 1999-01-27
NUMBER OF SEQ ID NOS: 485
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 375
LENGTH: 372
TYPE: PRT
ORGANISM: E. coli
US-09-912-020-375

Query Match 55.0%; Score 33; DB 10; Length 372;
Best Local Similarity 63.6%; Pred. No. 91;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 2 AEGIDTQRFSG 12
||| | | |
Db 250 AEGIDNVHFLG 260

RESULT 12
US-09-925-300-1674
Sequence 1674, Application US/09925300
Patent No. US20020151681A1
GENERAL INFORMATION:
APPLICANT: Craig Rosen,
APPLICANT: Steve Ruben
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies

FILE REFERENCE: PA101
CURRENT APPLICATION NUMBER: US/09/925,300
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05988
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1890
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1674
LENGTH: 375
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (338)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (340)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (356)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (372)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1674

Query Match 55.0%; Score 33; DB 10; Length 375;
Best Local Similarity 50.0%; Pred. No. 91;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 AAEGLDTQRFSG 12
|| | | | |
Db 62 AANGNDSKRFKG 73

RESULT 13
US-09-895-828-452
Sequence 452, Application US/09895828
Patent No. US20020099012A1
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: McNeill, Patricia D.
APPLICANT: Matanabe, Yoshihiro
APPLICANT: Carter, Darrick
APPLICANT: Henderson, Robert A.
APPLICANT: Kalos, Michael D.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 210121.539
CURRENT APPLICATION NUMBER: US/09/895,828
CURRENT FILING DATE: 2001-06-28
NUMBER OF SEQ ID NOS: 473
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 452
LENGTH: 550
TYPE: PRT
ORGANISM: Homo sapiens
US-09-895-828-452

Query Match 55.0%; Score 33; DB 10; Length 550;
Best Local Similarity 50.0%; Pred. No. 14+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 AAEGLDTQRFSG 12
|| | | | |
Db 38 AANGNDSKRFKG 49

RESULT 14
US-09-815-242-13814
Sequence 13814, Application US/09815242

```

; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 13814
; TYPE: PRT
; LENGTH: 635
; ORGANISM: Salmonella typhi
US-09-815-242-13814

Query Match          55.0%; Score 33; DB 10; Length 635;
Best Local Similarity 85.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 DMOFSG 12
DB 427 ETOFSG 433

RESULT 15
US-09-738-626-3675
; Sequence 3675, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: YATEISHI, NAKO
; APPLICANT: SENO, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0

```

```

; SEQ ID NO 3675
; LENGTH: 689
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-3675

Query Match          55.0%; Score 33; DB 9; Length 689;
Best Local Similarity 58.3%; Pred. No. 1.8e+02;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAEGLDIOFSG 12
DB 608 AEEGLAEQEFNG 619

```

Search completed: April 15, 2003, 10:19:25
Job time : 40 secs

GenCore version 5.1.4_P5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 15, 2003, 10:11:09 ; Search time 29 Seconds
(without alignments)
12.175 Million cell updates/sec

Title: US-09-521-527C-1
Perfect score: 60
Sequence: 1 AAEGLDYGRFSG 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCrus.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	102	3	US-08-466-368-8
2	60	100.0	114	1	US-08-372-952-1
3	60	100.0	114	4	US-08-875-309-1
4	60	100.0	114	5	PCR-US96-00310-1
5	60	100.0	235	1	US-07-940-605A-12
6	60	100.0	235	2	US-08-690-096-12
7	60	100.0	273	2	US-08-403-853-12
8	60	100.0	376	3	US-08-751-512-8
9	44	73.3	15	4	US-08-875-309-15
10	44	73.3	16	4	US-08-875-309-16
11	40	66.7	8	1	US-08-372-952-3
12	40	66.7	8	4	US-08-875-309-3
13	40	66.7	8	5	PCR-US96-00310-3
14	40	66.7	10	1	US-08-372-952-6
15	40	66.7	10	4	US-08-875-309-6
16	40	66.7	10	5	PCR-US96-00310-6
17	36	60.0	553	3	US-08-630-172-13
18	36	60.0	553	4	US-09-375-419-13
19	35	58.3	78	2	US-08-245-511-39
20	35	58.3	78	2	US-08-600-993A-39
21	34	56.7	751	4	US-09-036-987A-24
22	34	56.7	751	4	US-09-370-700-24
23	32	53.3	513	2	US-09-122-230-7
24	31.5	52.5	107	2	US-08-888-366-14
25	31.5	52.5	107	2	US-08-888-366-20
26	31.5	52.5	107	2	US-08-888-366-26
27	31.5	52.5	107	3	US-08-397-411-2

28	31	51.7	79	2	US-08-469-537A-49	Sequence 49, Appl
29	31	51.7	148	4	US-09-134-001C-3640	Sequence 3640, Ap
30	31	51.7	256	2	US-08-469-537A-58	Sequence 58, Appl
31	31	51.7	266	2	US-07-857-224B-75	Sequence 75, Appl
32	31	51.7	282	2	US-08-701-191A-19	Sequence 19, Appl
33	31	51.7	287	1	US-07-952-817-24	Sequence 24, Appl
34	31	51.7	341	4	US-09-149-624-2	Sequence 2, Appl1
35	31	51.7	425	4	US-08-462-467B-16	Sequence 16, Appl
36	31	51.7	432	3	US-08-985-908-31	Sequence 31, Appl
37	31	51.7	476	4	US-09-306-593-9	Sequence 9, Appl1
38	31	51.7	649	4	US-08-462-467B-20	Sequence 20, Appl
39	31	51.7	717	6	5262177-5	Patent No. 5262177
40	31	51.7	719	2	US-08-520-933-3	Sequence 3, Appl1
41	31	51.7	719	4	US-09-285-040-3	Sequence 31, Appl
42	31	51.7	738	6	5262177-2	Patent No. 5262177
43	31	51.7	1495	4	US-08-462-467B-12	Sequence 12, Appl
44	31	51.7	3218	1	US-08-764-100-27	Sequence 27, Appl
45	30	50.0	14	4	US-09-242-435-17	Sequence 17, Appl

ALIGNMENTS

RESULT 1
US-08-466-368-8
; Sequence 8, Application US/08466368
; Patent No. 6093539
; GENERAL INFORMATION:
; APPLICANT: Maddon, Paul J.
; APPLICANT: Littman, Dan R.
; APPLICANT: Chess, Leonard
; APPLICANT: Axel, Richard
; APPLICANT: Weiss, Robin
; APPLICANT: McDougall, J. S.
; TITLE OF INVENTION: DNA ENCODING THE T CELL SURFACE PROTEIN
; TITLE OF INVENTION: T4 AND USE OF FRAGMENTS OF T4 IN THE TREATMENT OF AIDS
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,368
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 24577-E1-B/JFW/AMC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 102 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: YES
; ANTI-SENSE: YES
; FEATURE:
; NAME/KEY: Active-site
; LOCATION: 1..102
; ;
; US-08-466-368-8

Query Match 100.0%; Score 60; DB 3; Length 102;
Best Local Similarity 100.0%; Pred. No. 0.0002;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAEGLDTRFSG 12
|||||
DB 59 AAEGLDTRFSG 70

RESULT 2

US-08-372-952-1
; Sequence 1, Application US/08372952
; Patent No. 5645837
; GENERAL INFORMATION:
; APPLICANT: Jameson, Bradford A.
; APPLICANT: Choksi, Swati
; APPLICANT: Korngold, Robert
; TITLE OF INVENTION: CD8 Antagonists
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
; ADDRESSEE: No. 5645837ris
; STREET: One Liberty Place, 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/372,952
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-1440
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 114 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-372-952-1

Query Match 100.0%; Score 60; DB 1; Length 114;
Best Local Similarity 100.0%; Pred. No. 0.0002;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAEGLDTRFSG 12
|||||
DB 59 AAEGLDTRFSG 70

RESULT 3

US-08-875-309-1
; Sequence 1, Application US/08875309
; Patent No. 6180600
; GENERAL INFORMATION:
; APPLICANT: Jameson, Bradford A.
; APPLICANT: Choksi, Swati
; APPLICANT: Korngold, Robert

APPLICANT: Huang, Ziwel
; TITLE OF INVENTION: CD8 Antagonists
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 6180600ris
; STREET: One Liberty Place, 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/875,309
; FILING DATE: 26-NOV-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/00310
; FILING DATE: 17-JAN-1996
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/372,952
; FILING DATE: 17-JAN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-1772
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 114 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-875-309-1

Query Match 100.0%; Score 60; DB 4; Length 114;
Best Local Similarity 100.0%; Pred. No. 0.0002;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAEGLDTRFSG 12
|||||
DB 59 AAEGLDTRFSG 70

RESULT 4

PCT-US96-00310-1
; Sequence 1, Application PC/TUS9600310
; GENERAL INFORMATION:
; APPLICANT: Jameson, Bradford A.
; APPLICANT: Choksi, Swati
; APPLICANT: Korngold, Robert
; APPLICANT: Huang, Ziwel
; TITLE OF INVENTION: CD8 Antagonists
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & Norris
; STREET: One Liberty Place, 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordperfect

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US96/00310
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/372,952
;; FILING DATE: 17-JAN-1995
;; CLASSIFICATION:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Deluca, Mark
;; REGISTRATION NUMBER: 33,229
;; REFERENCE/DOCKET NUMBER: TJJU-1752
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 215-568-3100
;; TELEFAX: 215-568-3439
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 114 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; PCT-US96-00310-1

Query Match 100.0%; Score 60; DB 5; Length 114;
Best Local Similarity 100.0%; Pred. No. 0.00022;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAEGDITORFSG 12
DB 59 AAEGDITORFSG 70

RESULT 5
US-07-940-605A-12
; Sequence 12, Application US/07940605A
; Patent No. 5540926
; GENERAL INFORMATION:
; APPLICANT: ARUEFO, ALEJANDRO
; APPLICANT: HOLLENBAUGH, DIANE
; APPLICANT: LEDBETTER, JEFFREY A.
; TITLE OF INVENTION: SOLUBLE LIGANDS FOR CD40
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/940,605A
; FILING DATE: 04-SEP-1992
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 5624-184
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 235 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-07-940-605A-12

Query Match 100.0%; Score 60; DB 1; Length 235;
Best Local Similarity 100.0%; Pred. No. 0.0005;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAEGDITORFSG 12
DB 80 AAEGDITORFSG 91

RESULT 6
US-08-690-096-12
; Sequence 12, Application US/08690096
; Patent No. 5945513
; GENERAL INFORMATION:
; APPLICANT: ARUEFO, ALEJANDRO
; APPLICANT: HOLLENBAUGH, DIANE
; APPLICANT: LEDBETTER, JEFFREY A.
; TITLE OF INVENTION: SOLUBLE LIGANDS FOR CD40
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/690,096
; FILING DATE: 31-JUL-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/940,605
; FILING DATE: 04-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 5624-184
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 235 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-690-096-12

Query Match 100.0%; Score 60; DB 2; Length 235;
Best Local Similarity 100.0%; Pred. No. 0.0005;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAEGDITORFSG 12
DB 80 AAEGDITORFSG 91

RESULT 7
US-08-403-853-12
; Sequence 12, Application US/08403853
; Patent No. 5844094
; GENERAL INFORMATION:
; APPLICANT: HUDSON, Peter J.
; APPLICANT: LAH, Maria
; APPLICANT: KORT, Alex A.
; APPLICANT: IRVING, Robert A.

APPLICANT: ATWELL, John L.
APPLICANT: MALBY, Robyn L.
APPLICANT: POWER, Barbara E.
APPLICANT: COLMAN, Peter M.
TITLE OF INVENTION: TARGET BINDING POLYPEPTIDE
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/403,853
FILING DATE: 30-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/AU93/00491
FILING DATE: 24-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PL 4973
FILING DATE: 25-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: BERT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16786/189/CHAC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 273 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-403-853-12

Query Match 100.0%; Score 60; DB 2; Length 273;
Best Local Similarity 100.0%; Pred. No. 0.00058;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAEGLDTRFSG 12
DB 81 AAEGLDTRFSG 92

RESULT 8
US-08-751-512-8
Sequence 8, Application US/08751512
Patent No. 6001962
GENERAL INFORMATION:
APPLICANT: Ramer, J. Kevin
APPLICANT: Williams, Lewis T.
TITLE OF INVENTION: Modified FAS Ligands
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: TOWNSEND and TOWNSEND and CREM LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: IBM PC compatible
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/751,512
FILING DATE: 15-NOV-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Murphy, Matthew B.
REGISTRATION NUMBER: 39,787
REFERENCE/DOCKET NUMBER: 02307K-07100US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-576-0200
TELEFAX: 415-576-0300
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 376 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-751-512-8

Query Match 100.0%; Score 60; DB 3; Length 376;
Best Local Similarity 100.0%; Pred. No. 0.00083;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAEGLDTRFSG 12
DB 80 AAEGLDTRFSG 91

RESULT 9
US-08-875-309-15
Sequence 15, Application US/08875309
Patent No. 6180600
GENERAL INFORMATION:
APPLICANT: Jameson, Bradford A.
APPLICANT: Choksi, Swati
APPLICANT: Krongold, Robert
APPLICANT: Huang, Zilwei
TITLE OF INVENTION: CD8 Antagonists
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz MacKiewicz & No. 6180600r1s
STREET: One Liberty Place, 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/875,309
FILING DATE: 26-NOV-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/00310
FILING DATE: 17-JAN-1996
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/372,952
FILING DATE: 17-JAN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Deluca, Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: TJU-1772
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:

LENGTH: 15 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-875-309-15

Query Match 73.3%; Score 44; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.028;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAEGLDTR 9
|||||
Db 7 AAEGLDTR 15

RESULT 10
US-08-875-309-16
Sequence 16, Application US/08875309
Patent No. 6180600
GENERAL INFORMATION:
APPLICANT: Jameson, Bradford A.
APPLICANT: Choksi, Swati
APPLICANT: Korngold, Robert
APPLICANT: Huang, Ziwei
TITLE OF INVENTION: CD8 Antagonists
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 6180600r1s
STREET: One Liberty Place, 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/875,309
FILING DATE: 26-NOV-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/00310
FILING DATE: 17-JAN-1996
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/372,952
FILING DATE: 17-JAN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Deluca, Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: TJU-1772
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-875-309-16

Query Match 73.3%; Score 44; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAEGLDTR 9
|||||
Db 8 AAEGLDTR 16

RESULT 11
US-08-372-952-3
Sequence 3, Application US/08372952
Patent No. 5645837
GENERAL INFORMATION:
APPLICANT: Jameson, Bradford A.
APPLICANT: Choksi, Swati
APPLICANT: Korngold, Robert
TITLE OF INVENTION: CD8 Antagonists
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
ADDRESSEE: No. 5645837r1s
STREET: One Liberty Place, 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/372,952
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Deluca, Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: TJU-1440
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-372-952-3

Query Match 66.7%; Score 40; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.9e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 AEGIDTR 9
|||||
Db 1 AEGIDTR 8

RESULT 12
US-08-875-309-3
Sequence 3, Application US/08875309
Patent No. 6180600
GENERAL INFORMATION:
APPLICANT: Jameson, Bradford A.
APPLICANT: Choksi, Swati
APPLICANT: Korngold, Robert
APPLICANT: Huang, Ziwei
TITLE OF INVENTION: CD8 Antagonists
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 6180600r1s
STREET: One Liberty Place, 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA

```

;
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/875,309
; FILING DATE: 26-NOV-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/00310
; FILING DATE: 17-JAN-1996
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/372,952
; FILING DATE: 17-JAN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-1772
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-875-309-3

Query Match
Best Local Similarity 66.7%; Score 40; DB 4; Length 8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 AEGDITOR 9
DB 1 AEGDITOR 8

RESULT 13
PCT-US96-00310-3
; Sequence 3, Application PC/TUS9600310
; GENERAL INFORMATION:
; APPLICANT: Jameson, Bradford A.
; APPLICANT: Choksi, Swati
; APPLICANT: Korgold, Robert
; APPLICANT: Huang, Ziwel
; TITLE OF INVENTION: CD8 Antagonists
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & Norris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/00310
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/372,952
; FILING DATE: 17-JAN-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
```

```

;
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-1752
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US96-00310-3
```

```

Query Match
Best Local Similarity 66.7%; Score 40; DB 5; Length 8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 AEGDITOR 9
DB 1 AEGDITOR 8
```

```

RESULT 14
US-08-372-952-6
; Sequence 6, Application US/08372952
; Patent No. 5645837
; GENERAL INFORMATION:
; APPLICANT: Jameson, Bradford A.
; APPLICANT: Choksi, Swati
; APPLICANT: Korgold, Robert
; TITLE OF INVENTION: CD8 Antagonists
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/372,952
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-1440
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-372-952-6

Query Match
Best Local Similarity 66.7%; Score 40; DB 1; Length 10;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 2 AEGLDTOR 9
|||||||
Db 2 AEGLDTOR 9

RESULT 15
US-08-875-309-6

; Sequence 6, Application US/08875309
; Patent No. 6180600
; GENERAL INFORMATION:
; APPLICANT: Jameson, Bradford A.
; APPLICANT: Choksi, Swati
; APPLICANT: Korngold, Robert
; APPLICANT: Huang, Ziwei
; TITLE OF INVENTION: CD8 Antagonists
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 6180600ris
; STREET: One Liberty Place, 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/875,309
; FILING DATE: 26-NOV-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/00310
; FILING DATE: 17-JAN-1996
; CLASSIFICATION: 424
; APPLICATION NUMBER: US 08/372,952
; FILING DATE: 17-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-1772
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ. ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: circular
; MOLECULE TYPE: protein
; US-08-875-309-6

Query Match 66.7%; Score 40; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AEGLDTOR 9
|||||||
Db 2 AEGLDTOR 9

Search completed: April 15, 2003, 10:13:56
Job time : 31 secs

GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 15, 2003, 10:18:44 ; Search time 14 Seconds

(without alignments)
82.401 Million cell updates/sec

Title: US-09-521-527C-1

Perfect score: 60

Sequence: 1 AAEGLDTRFSG 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 segs, 96134422 residues

Total number of hits satisfying chosen parameters: 1580

Minimum DB seq length: 0
Maximum DB seq length: 12

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR73:*

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	28	46.7	11	2 A35594	buccalin - Callifor
2	23	38.3	9	2 F41978	callifmrfamide 6 -
3	22	36.7	9	2 E41978	callifmrfamide 5 -
4	22	36.7	9	2 G41978	callifmrfamide 7 -
5	22	36.7	12	2 G49410	t-complex polypept
6	22	36.7	12	2 A33099	163k exoantigen -
7	21	35.0	10	2 PNO136	pepsin (EC 3.4.23.
8	21	35.0	12	2 P41481	T-cell receptor be
9	20	33.3	12	2 S07436	tachykinin - Afric
10	19	31.7	8	2 S13661	polygalacturonase
11	19	31.7	9	2 B33098	231k exoantigen -
12	18	30.0	9	2 D57444	neuropeptide Grb-A
13	18	30.0	10	2 P70251	Ig heavy chain CRD
14	18	30.0	12	2 PH0920	T-cell receptor be
15	17	28.3	7	2 P70529	T-cell receptor be
16	17	28.3	10	2 S70721	heat shock protein
17	17	28.3	11	2 S07201	physalaemin - frog
18	17	28.3	11	2 S19775	wound-induced prot
19	16	26.7	7	2 S09066	globulin IV alpha
20	16	26.7	8	2 B24749	neuropeptide B - b
21	16	26.7	8	2 S71919	alcohol dehydrogen
22	16	26.7	9	2 A61230	caesquestrin, car
23	16	26.7	9	2 C24180	fibrinogen beta ch
24	16	26.7	10	2 S33844	alpha-2-macroglobu
25	16	26.7	10	2 PH0946	T-cell receptor be
26	16	26.7	11	1 SPHO	substance P - hors
27	16	26.7	11	1 A60654	substance P - quin
28	16	26.7	11	1 JN0023	substance P - chic
29	16	26.7	11	2 S23306	substance P - Atla

30	16	26.7	11	2 A61512	variant surface gl
31	16	26.7	11	2 PT0287	Ig heavy chain CRD
32	16	26.7	11	2 PT0217	T-cell receptor be
33	16	26.7	11	2 A48973	glucosylase A1 (E
34	16	26.7	11	2 H84082	hypothetical prote
35	16	26.7	12	2 S01749	collagen alpha 1(I
36	16	26.7	12	2 S25039	Ig heavy chain V r
37	16	26.7	12	2 PH0802	T-cell receptor al
38	16	26.7	12	2 PH0790	hexokinase (EC 2.7
39	16	26.7	12	2 B32521	T-cell receptor al
40	16	26.7	12	4 PC2123	aminotransferase C
41	15	25.0	6	2 PT0718	T-cell receptor be
42	15	25.0	7	2 PT0663	T-cell receptor be
43	15	25.0	8	2 H41978	callifmrfamide 8 -
44	15	25.0	8	2 PT0554	T-cell receptor be
45	15	25.0	9	2 A61357	phyllocerulolein -

ALIGNMENTS

RESULT 1
A35594
buccalin - California sea hare
C:Species: Aplysia californica (California sea hare)
C:Date: 14-Sep-1990 #sequence_revision 14-Sep-1990 #text_change 24-Jun-1993
C:Accession: A35594
R: Cropper, E.C.; Miller, M.W.; Tenenbaum, R.; Kolks, M.A.G.; Kupfermann, I.; Weiss, K
Proc. Natl. Acad. Sci. U.S.A. 85, 6177-6181, 1988
A:Title: Structure and action of buccalin: a modulatory neuropeptide localized to an
A:Reference number: A35594; PMID:8320404; PMID:3413086
A:Accession: A35594
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-11 <CRO>

Query Match 46.7%; Score 28; DB 2; Length 11;
Best Local Similarity 55.6%; Pred. 48;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 GLDTRFSG 12
DB 1 GMDSLAFSG 9

RESULT 2
F41978
callifmrfamide 6 - bluebottle fly (Calliphora vomitoria)
C:Species: Calliphora vomitoria
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 17-Mar-1999
C:Accession: F41978
R: Duve, H.; Johnsen, A.H.; Sewell, J.C.; Scott, A.G.; Orchard, I.; Rehfeld, J.F.; Tho
Proc. Natl. Acad. Sci. U.S.A. 89, 2326-2330, 1992
A:Title: Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH-2 neuropeptides (d
A:Reference number: A41978; PMID:92196111; PMID:1549595
A:Accession: F41978
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-9 <DOV>
C:Keywords: amidated carboxyl end; neuropeptide
F:9/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match 38.3%; Score 23; DB 2; Length 9;
Best Local Similarity 55.6%; Pred. No. 2.8e+05;
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAEGLDTRF 10
DB 1 ASGQDFMRF 9

RESULT 3
E41978

callIFMRamide 5 - bluebottle fly (Calliphora vomitoria)

C:Species: Calliphora vomitoria

C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 17-Mar-1999

C:Accession: E41978

R:Dive, H.; Johnson, A.H.; Sewell, J.C.; Scott, A.G.; Orchard, I.; Rehfeld, J.F.; Thorpe

Proc. Natl. Acad. Sci. U.S.A. 89, 2326-2330, 1992

A:Title: Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH-2 neuropeptides (des

A:Reference number: A41978; MUID:92196111; PMID:1549595

C:Accession: E41978

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-9 <DUV>

C:Keywords: amidated carboxyl end; neuropeptide

F:9/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match

Best Local Similarity 36.7%; Score 22; DB 2; Length 9;

Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 2 AEGDTPORF 10

DB 1 AFGDFMRF 9

RESULT 4

callIFMRamide 7 - bluebottle fly (Calliphora vomitoria)

C:Species: Calliphora vomitoria

C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 17-Mar-1999

C:Accession: G41978

R:Dive, H.; Johnson, A.H.; Sewell, J.C.; Scott, A.G.; Orchard, I.; Rehfeld, J.F.; Thorpe

Proc. Natl. Acad. Sci. U.S.A. 89, 2326-2330, 1992

A:Title: Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH-2 neuropeptides (des

A:Reference number: A41978; MUID:92196111; PMID:1549595

C:Accession: G41978

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-9 <DUV>

C:Keywords: amidated carboxyl end; neuropeptide

F:9/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match

Best Local Similarity 36.7%; Score 22; DB 2; Length 9;

Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 2 AEGDTPORF 10

DB 1 AXGDFMRF 9

RESULT 5

t-complex polypeptide 1 homolog (peak 6b fraction) - rabbit (fragment)

N:Alternate names: Chaperonin homolog (peak 6b)

C:Species: Oryctolagus cuniculus (domestic rabbit)

C>Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 04-Sep-1998

C:Accession: G49410

R:Rommelaere, H.; Van Troys, M.; Gao, Y.; Melki, R.; Cowan, N.J.; Vandekerckhove, J.; An

Proc. Natl. Acad. Sci. U.S.A. 90, 11975-11979, 1993

A:Title: Eukaryotic cytosolic chaperonin contains t-complex polypeptide 1 and seven rela

A:Reference number: A49410; MUID:94089752; PMID:7903455

C:Accession: G49410

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-12 <ROM>

A:Experimental source: reticulocyte

C:Superfamily: molecular chaperone t-complex-type

Query Match

Best Local Similarity 36.7%; Score 22; DB 2; Length 12;

Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 AAEGLDTPOR 9

DB 4 ANTCMDTRDK 12

RESULT 6

A33099

163k exantigen - malaria parasite (Plasmodium falciparum) (fragments)

C:Species: Plasmodium falciparum

C>Date: 24-Aug-1990 #sequence_revision 24-Aug-1990 #text_change 09-Jun-2000

C:Accession: A33099

R:Nichols, J.H.; Hager, L.P.

submitted to the Protein Sequence Database, May 1990

A:Reference number: A33098

A:Accession: A33099

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-12 <NIC>

Query Match

Best Local Similarity 36.7%; Score 22; DB 2; Length 12;

Matches 4; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 4 GIDTPQRESG 12

DB 1 GVDTPQEVGG 9

RESULT 7

PN0136

pepsin (EC 3.4.23.-) 4 - horse (fragment)

N:Alternate names: pepsin (PI 2.6)

C:Species: Equus caballus (domestic horse)

C>Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 07-May-1999

C:Accession: PN0136

R:Gonchar, M.V.; Lavrenova, G.I.; Rudenskaya, G.N.; Galda, A.V.; Stepanov, V.M.

Biochimia 49, 1026-1037, 1984

A:Title: Multiple forms of horse pepsin.

A:Reference number: PN0133; MUID:84281135; PMID:6432065

A:Accession: PN0136

A:Molecule type: protein

A:Residues: 1-10 <GON>

A:Note: article in Russian with English abstract

C:Superfamily: pepsin

C:Keywords: aspartic proteinase; hydrolase; protein digestion

Query Match

Best Local Similarity 35.0%; Score 21; DB 2; Length 10;

Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 AAEGLD 6

DB 2 ATBGL 7

RESULT 8

PH1481

T-cell receptor beta chain (clone A24/PEG2) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 11-Apr-1995

C:Accession: PH1481

R:Casanova, J.L.; Martinou, F.; Gounluer, H.; Barra, C.; Pannetier, C.; Regnault, A.;

J. Exp. Med. 177, 811-820, 1993

A:Title: T cell receptor selection by and recognition of two class I major histocompa

A:Reference number: PH1430; MUID:93171821; PMID:8436911

A:Accession: PH1481

A:Molecule type: mRNA

A:Residues: 1-12 <CAS>

A:Experimental source: cytoytic T-lymphocyte

A:Note: the authors translated the codon GGT for residue 2 as Ala and GCT for residue

C:Superfamily: immunoglobulin homology

C:Keywords: receptor; T-cell

Query Match

Best Local Similarity 35.0%; Score 21; DB 2; Length 12;

Best Local Similarity 50.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 2; Mismatches 2; Indels 2; Gaps 1;

OY 1 AAEGDTPRFSG 12
||| :||
DB 3 AAMG--AEQFPG 12

RESULT 9

S07436
tachykinin - African tree frog (*Kassina maculata*)

N:Alternate names: hylambatin

C:Species: *Kassina maculata*

C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 02-Sep-2000

C:Accession: S07436

R:Yasuhara, T.; Nakajima, T.; Erspamer, G.F.; Erspamer, V.

Biomed. Res. 2, 613-617, 1981

A>Title: New tachykinins, Glu2, Pro5-kassinin (*hylambates-kassinin*) and *hylambatin*, in t

A:Reference number: S07436

A:Accession: S07436

A:Molecule type: protein

A:Residues: 1-12 <YAS>

A:Experimental source: skin

A>Note: the source is designated as *hylambates maculatus*

C:Superfamily: unassigned animal peptides

C:Keywords: amidated carboxyl end; neuropeptide; tachykinin

F:12/Modified site: amidated carboxyl end (Met) #status predicted

Query Match 33.3%; Score 20; DB 2; Length 12;
Best Local Similarity 57.1%; Pred. No. 1.8e+03;

Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 6 DTPQRFSG 12
||| :||
DB 4 DPDFRYG 10

RESULT 10

S13661
polygalacturonase (EC 3.2.1.15) isoform PG2 - fungus (*Sclerotinia sclerotiorum*) (fragment

N:Alternate names: endopolygalacturonase; pectin depolymerase; pectinase

C:Species: *Sclerotinia sclerotiorum*

C>Date: 19-Mar-1997 #sequence_revision 21-Nov-1998 #text_change 21-Nov-1998

C:Accession: S13661

R:Waksman, G.; Keon, J.P.R.; Turner, G.

Biochim. Biophys. Acta 1073, 43-48, 1991

A>Title: Purification and characterization of two endopolygalacturonases from *Sclerotinia*

A:Reference number: S13661; MUID:91120822; PMID:1991145

A:Accession: S13661

A:Molecule type: protein

A:Residues: 1-8 <MAK>

A:Function:

A:Description: involved in pectin degradation

C:Keywords: glycosidase; hydrolase

Query Match 31.7%; Score 19; DB 2; Length 8;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 7 TQRFSG 12
||| :||
DB 3 TDTFSG 8

RESULT 11

B33098
23k exoantigen - malaria parasite (*Plasmodium falciparum*) (fragments)

C:Species: *Plasmodium falciparum*

C>Date: 24-Aug-1990 #sequence_revision 24-Aug-1990 #text_change 09-Jun-2000

C:Accession: B33098

R:Nichols, J.H.; Hager, L.P.

submitted to the Protein Sequence Database, May 1990

A:Reference number: A33098

A:Accession: B33098
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-9 <NIC>

Query Match 31.7%; Score 19; DB 2; Length 9;
Best Local Similarity 44.4%; Pred. No. 2.8e+05;
Matches 4; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 4 GLDTPRFSG 12
||| :||
DB 1 GLDTAXIXG 9

RESULT 12

D57444
neuropeptide Grb-AsT B4 - two-spotted cricket

C:Species: *Gryllus bimaculatus* (two-spotted cricket)

C>Date: 26-Jan-1996 #sequence_revision 26-Jan-1996 #text_change 26-Jan-1996

C:Accession: D57444

R:Lorenz, M.W.; Kellner, R.; Hoffmann, K.H.

J. Biol. Chem. 270, 21103-21108, 1995

A>Title: A family of neuropeptides that inhibit juvenile hormone biosynthesis in the

A:Reference number: A57444; MUID:95403341; PMID:7673141

A:Accession: D57444

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-9 <LOR>

Query Match 30.0%; Score 18; DB 2; Length 9;
Best Local Similarity 60.0%; Pred. No. 2.8e+05;

Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 8 ORFSG 12
||| :||
DB 3 ERFHG 7

RESULT 13

PT0251

Ig heavy chain CDR3 region (clone 2-109C) - human (fragment)

C:Species: *Homo sapiens* (man)

C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996

C:Accession: PT0251

R:Yamada, M.; Masserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.

J. Exp. Med. 173, 395-407, 1991

A>Title: Preferential utilization of specific immunoglobulin heavy chain diversity an

A:Reference number: PT0222; MUID:91108337; PMID:1899102

A:Accession: PT0251

A:Molecule type: DNA

A:Residues: 1-10 <YAM>

A:Experimental source: B lymphocyte

C:Keywords: heterotetramer; immunoglobulin

Query Match 30.0%; Score 18; DB 2; Length 10;
Best Local Similarity 60.0%; Pred. No. 3.7e+03;

Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 7 TQRFSG 11
||| :||
DB 1 SORFT 5

RESULT 14

PH0920

T-cell receptor beta chain V-D-J region (isolate 6) - rat (fragment)

C:Species: *Rattus norvegicus* (Norway rat)

C>Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 30-May-1997

C:Accession: PH0920

R:Gold, D.P.; Offner, H.; Sun, D.; Wiley, S.; Vandenbark, A.A.; Wilson, D.B.

J. Exp. Med. 174, 1467-1476, 1991

A>Title: Analysis of T cell receptor beta chains in Lewis rats with experimental alle

A:Reference number: PH0991; MUID:92078857; PMID:1836012

A:Accession: PH0920
A:Molecule type: mRNA
A:Residues: 1-12 <GOL>
A:Experimental source: concanavalin A-activated lymphoblast
A>Note: the authors translated the codon CAG for residue 12 as Glu
C:Keywords: T-cell receptor

Query Match 30.0%; Score 18; DB 2; Length 12;
Best Local Similarity 42.9%; Pred. No. 4.4e+03;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 AAEGLD 7
:|:|:|
DB 4 SAQGANT 10

RESULT 15

PT0529
T-cell receptor beta chain V-D-J region (100-4K) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0529
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:9127601; PMID:1711558
A:Accession: PT0529
A>Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-7 <FE>
A:Experimental source: adult thymus, strain BALB/c
C:Keywords: T-cell receptor

Query Match 28.3%; Score 17; DB 2; Length 7;
Best Local Similarity 60.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAEG 5
::|||
DB 2 SSEGL 6

Search completed: April 15, 2003, 10:21:19
Job time : 15 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 15, 2003, 10:14:54 ; Search time 11 Seconds

(without alignments)
45.247 Million cell updates/sec

Title: US-09-521-527C-1

Perfect score: 60

Sequence: 1 AAEGLDTQRFSG 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 467

Minimum DB seq length: 0
Maximum DB seq length: 12

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	23	38.3	9	1	FAR6_CALVO
2	22	36.7	9	1	FAR5_CALVO
3	22	36.7	9	1	FAR7_CALVO
4	20	33.3	12	1	TKN2_KASMA
5	18	30.0	11	1	TKC2_CALVO
6	17	28.3	11	1	TKNA_RANRI
7	17	28.3	11	1	TKN_PHYFU
8	16	26.7	8	1	NPB_BOVIN
9	16	26.7	9	1	FIBB_MACFU
10	16	26.7	10	1	URAI_HUMAN
11	16	26.7	11	1	TKNA_CHICK
12	16	26.7	11	1	TKNA_HUMAN
13	16	26.7	11	1	TKNA_HORSE
14	15	25.0	8	1	ALIS_CYDPO
15	15	25.0	8	1	FAR8_CALVO
16	15	25.0	9	1	FAR1_CALVO
17	15	25.0	9	1	FAR2_CALVO
18	15	25.0	9	1	FAR3_CALVO
19	15	25.0	9	1	FAR4_CALVO
20	15	25.0	9	1	FAR5_CALVO
21	15	25.0	9	1	FIBB_PAPNA
22	15	25.0	9	1	FIBB_PAPNA
23	15	25.0	9	1	FIBB_PAPNA
24	15	25.0	10	1	AMPN_HELAM
25	15	25.0	10	1	COXO_THIOB
26	15	25.0	10	1	TKN_PHYBI
27	15	25.0	11	1	FAR9_CALVO
28	15	25.0	11	1	TKN2_UPERU
29	15	25.0	11	1	TKNA_RANCA
30	15	25.0	12	1	SOIS_BACSU
31	15	25.0	12	1	UR2B_CATCO
32	14	23.3	8	1	CUP_THICU
33	14	23.3	10	1	PKCK_FASHE

34	14	23.3	10	1	TKNB_CHICK	P18611
35	14	23.3	11	1	PVK1_PERAM	P41837
36	14	23.3	11	1	RR2_CONAM	P42341
37	14	23.3	11	1	TKNI_UPEIN	P82026
38	14	23.3	11	1	TKNI_UPERU	P82026
39	13	22.5	10	1	FARC_CALVO	P41867
40	13	21.7	8	1	COXG_RAT	P80430
41	13	21.7	9	1	LMT2_LOCOMI	P22396
42	13	21.7	9	1	CCAP_CARMA	P38556
43	13	21.7	10	1	CA12_LITCI	P82086
44	13	21.7	10	1	CAER_LITXA	P56264
45	13	21.7	10	1	VEG6_BACSU	P80659

ALIGNMENTS

```

RESULT 1
ID FAR6_CALVO STANDARD; PRT; 9 AA.
AC P41861;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE CalliFMRamide 6.
OS Calliphora vomitoria (Blue blowfly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Oestroidea; Calliphoridae; Calliphora.
OX NCBI_TaxID=27454;
RN [1]
RP SEQUENCE.
RC TISSUE=Thoracic ganglion;
RX MEDLINE=92196111; PubMed=1549595;
RA Duvé H., Johnsen A.H., Sewell J.C., Scott A.G., Orchard I.,
RA Rehfeld J.F., Thorpe A.;
RT "Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH2
RT neuropeptides (designated calliFMRamides) from the blowfly
RT Calliphora vomitoria.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:2326-2330(1992)
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRAMIDE RELATED PEPTIDE)
CC FAMILY.
CC DR PIR: F41978; F41978.
CC KW Neuropeptide; Amidation.
CC FT MOD_RES 9
CC SQ SEQUENCE 9 AA: 1058 MW: 96010699CAB6D865 CRC64:
Query Match 38.3%; Score 23; DB 1; Length 9;
Best Local Similarity 55.6%; Pred. No. 1.1e+05;
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 AAEGLDTQRF 10
Db 1 ASGQDFMRF 9
RESULT 2
ID FAR5_CALVO STANDARD; PRT; 9 AA.
AC P41860;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE CalliFMRamide 5.
OS Calliphora vomitoria (Blue blowfly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Oestroidea; Calliphoridae; Calliphora.
OX NCBI_TaxID=27454;
RN [1]
RP SEQUENCE.
RC TISSUE=Thoracic ganglion;
RX MEDLINE=92196111; PubMed=1549595;

```

RA Duvé H., Johnson A.H., Sewell J.C., Scott A.G., Orchard I.,
 RA Rehfeld J.F., Thorpe A.;
 RT "Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH2
 RT neuropeptides (designated calliphramide) from the blowfly
 RT Calliphora vomitoria";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:2326-2330(1992).
 CC -1- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
 CC FAMILY.
 CC DR PIR: E41978; E41978.
 CC KW Neuropeptide; Amidation.
 FT MOD_RES 9
 SQ SEQUENCE 9 AA; 1068 MW; 39D10699CAB6D867 CRC64;
 OY 2 AEGDITQRF 10
 DB 1 AFGDPMRF 9

Query Match 36.7%; Score 22; DB 1; Length 9;
 Best Local Similarity 55.6%; Pred. No. 1.1e+05;
 Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

RESULT 3
 FARP_CALVO STANDARD; PRT; 9 AA.
 AC P41862;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Calliphramide 7.
 OS Calliphora vomitoria (Blue blowfly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 CC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 CC Muscomorpha; Oestroidea; Calliporidae; Calliphora.
 OX NCBI_TaxID=27454;
 RN [1]
 RP SEQUENCE.
 RC TISSUE:Thoracic ganglion;
 RX MEDLINE=9219611; PubMed=1549595;
 RA Duvé H., Johnson A.H., Sewell J.C., Scott A.G., Orchard I.,
 RA Rehfeld J.F., Thorpe A.;
 RT "Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH2
 RT neuropeptides (designated calliphramide) from the blowfly
 RT Calliphora vomitoria";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:2326-2330(1992).
 CC -1- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
 CC FAMILY.
 CC DR PIR: G41978; G41978.
 CC KW Neuropeptide; Amidation.
 FT MOD_RES 9
 SQ SEQUENCE 9 AA; 1081 MW; E1D10699CAB6D86A CRC64;
 OY 2 AEGDITQRF 10
 DB 1 AXGDPMRF 9

Query Match 36.7%; Score 22; DB 1; Length 9;
 Best Local Similarity 55.6%; Pred. No. 1.1e+05;
 Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

RESULT 4
 TRN2_KASMA STANDARD; PRT; 12 AA.
 AC P08614;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Hylambatin.
 OS Kassina maculata (African rhacophorid frog) (Hylambates maculatus).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Hyperoliidae;
 CC Kassina.

OX NCBI_TaxID=8414;
 RN [1]
 RP SEQUENCE.
 RA Yasuhara T., Nakajima T., Erespaner G.F., Erespaner V.;
 RT "New tachykinins, Glu2, Pro5-kassinin (hylambates-kassinin) and
 RT hylambatin, in the skin of the African rhacophorid frog Hylambates
 RT maculatus";
 RL Biomed. Res. 2:613-617(1981).
 CC -1- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS.
 CC EXCITE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
 CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
 CC MUSCLES.
 CC -1- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
 CC DR PIR: S07436; S07436.
 CC DR Interpro: IPR003580; Proctachykinin.
 CC DR Interpro: IPR002040; Tachykinin.
 CC DR Pfam: PF02202; Tachykinin; 1.
 CC DR SMART: SM00203; TK; 1.
 CC DR PROSITE: PS00267; TACHYKININ; 1.
 CC KW Tachykinin; Neuropeptide; Amidation; Amphibian skin.
 FT MOD_RES 12
 SQ SEQUENCE 12 AA; 1441 MW; 3287CD2FDD040AB7 CRC64;
 OY 6 DTQRFSG 12
 DB 4 DPNRFYG 10

Query Match 33.3%; Score 20; DB 1; Length 12;
 Best Local Similarity 57.1%; Pred. No. 4.6e+02;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

RESULT 5
 TKC2_CALVO STANDARD; PRT; 11 AA.
 AC P41518;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Callitachykinin II.
 OS Calliphora vomitoria (Blue blowfly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 CC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 CC Muscomorpha; Oestroidea; Calliporidae; Calliphora.
 OX NCBI_TaxID=27454;
 RN [1]
 RP SEQUENCE, AND SYNTHESIS.
 RX MEDLINE=95075727; PubMed=7984492;
 RA Lundquist C.T., Clottens F.L., Holman G.M., Nichols R., Nachman R.J.,
 RA Messel D.R.;
 RT "Callitachykinin I and II, two novel myotropic peptides isolated from
 RT the blowfly, Calliphora vomitoria, that have resemblances to
 RT tachykinins";
 RL Peptides 15:761-768(1994).
 CC -1- FUNCTION: MYOACTIVE PEPTIDE.
 CC -1- SIMILARITY: SOME SIMILARITY TO TACHYKININS.
 CC KW Tachykinin; Neuropeptide; Amidation.
 FT MOD_RES 11
 SQ SEQUENCE 11 AA; 1103 MW; 15D7E3F9C9CD0444 CRC64;
 OY 4 GLDTQRFSG 12
 DB 1 GLGNMRFVG 9

Query Match 30.0%; Score 18; DB 1; Length 11;
 Best Local Similarity 44.4%; Pred. No. 1.1e+03;
 Matches 4; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

RESULT 6
 TRNA_RANRI STANDARD; PRT; 11 AA.
 ID TRNA_RANRI
 AC P29207;

```

DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Ranakidin (Substance-P-related peptide).
OS Rana ridibunda (Laughing frog) (Marsh frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Ranidae; Rana.
OX NCBI_TaxID=8406;
RN [1]
RP SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=92044543; PubMed=1658233;
RA O'Harte F., Burcher E., Lovas S., Smith D.D., Vaudry H., Conlon J.M.;
RT "Ranakidin: a novel NK1 tachykinin receptor agonist isolated with
RT neurokinin B from the brain of the frog Rana ridibunda.";
RL J. Neurochem. 57:2086-2091(1991).
CC -1- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
CC MUSCLES.
CC -1- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
DR InterPro: IPR003580; Protachykinin.
DR InterPro: IPR002040; Tachykinin.
DR Pfam: PF02202; Tachykinin; 1.
DR SMART: SM00203; TK; 1.
DR PROSITE: PS00267; TACHYKININ; 1.
KW Tachykinin; Neuropeptide; Amidation.
FT MOD_RES 11 11 AMIDATION.
SQ SEQUENCE 11 AA; 1352 MW; 3A2460CC59D40B07 CRC64;

Query Match 28.3%; Score 17; DB 1; Length 11;
Best Local Similarity 60.0%; Pred. No. 1.7e+03;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 8 QRFSG 12
Db 5 ERFYG 9

RESULT 7
TKN_PHYFU STANDARD; PRT; 11 AA.
AC P08615;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Physalaemin.
OS Physalaemus fuscumaculatus (Neotropical frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Leptodactylidae;
OC Physalaemus.
OX NCBI_TaxID=8378;
RN [1]
RP SEQUENCE.
RX MEDLINE=66076612; PubMed=5857249;
RA Erspamer V., Anastasi A., Bertaccini G., Cei J.M.;
RT "Structure and pharmacological actions of physalaemin, the main
RT active polypeptide of the skin of Physalaemus fuscumaculatus.";
RL Experientia 20:489-490(1964).
CC -1- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
CC MUSCLES.
CC -1- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
DR InterPro: IPR003580; Protachykinin.
DR InterPro: IPR002040; Tachykinin.
DR Pfam: PF02202; Tachykinin; 1.
DR SMART: SM00203; TK; 1.
DR PROSITE: PS00267; TACHYKININ; 1.
KW Tachykinin; Neuropeptide; Amidation; Amphibian skin.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 11 11 AMIDATION.

```

```

SQ SEQUENCE 11 AA; 1283 MW; 3293693E59C33457 CRC64;

Query Match 28.3%; Score 17; DB 1; Length 11;
Best Local Similarity 42.9%; Pred. No. 1.7e+03;
Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 6 DTORFSG 12
Db 3 DPNKFYG 9

RESULT 8
NPB_BOVIN STANDARD; PRT; 8 AA.
ID NPB_BOVIN
AC P15507;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-APR-1990 (Rel. 14, Last annotation update)
DE Neuropeptide B.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=86067985; PubMed=3865193;
RA Yang H.-Y.T., Pratta W., Majane E.A., Costa E.;
RT "Isolation, sequencing, synthesis, and pharmacological
RT characterization of two brain neuropeptides that modulate the action
RT of morphine.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:7757-7761(1985).
CC -1- FUNCTION: MODULATES THE ACTION OF MORPHINE.
DR PIR: B24749; B24749.
KW Neuropeptide; Amidation.
FT MOD_RES 5 8 TO NEUROPEPTIDE A (AA 5-8) (IDENTICAL).
FT MOD_RES 8 8 AMIDATION.
FT MOD_RES 8 8
SQ SEQUENCE 8 AA; 1082 MW; 87D416C776D9C729 CRC64;

Query Match 26.7%; Score 16; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 QRF 10
Db 6 QRF 8

RESULT 9
FIBB_MACFU STANDARD; PRT; 9 AA.
ID FIBB_MACFU
AC P19345;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).
OS Macaca fuscata fuscata (Japanese macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9543;
RN [1]
RP SEQUENCE.
RX MEDLINE=85289140; PubMed=3928610;
RA Nakamura S., Takenaka O., Takahashi K.;
RT "Fibrinopeptides A and B of Japanese monkey (Macaca fuscata) and
RT patas monkey (Erythrocebus patas): their amino acid sequences,
RT restricted mutations, and a molecular phylogeny for macaques,
RT guenons, and baboons.";
RL J. Biochem. 97:1487-1492(1985).
CC -1- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT

```

CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
 CC AGGREGATION.
 CC -1 SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
 CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
 CC -1 MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
 CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA 6 BETA
 CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
 CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
 DR PIR, C24180; C24180.
 DR InterPro: IPR002181; Fibrinogen.C.
 DR PROSITE: PS00514; FIBRIN_AG_C DOMAIN: PARTIAL.
 KW Blood coagulation; Plasma.
 FT PEPTIDE 1 9 FIBRINOPEPTIDE B.
 FT NON_TER 9 9
 SO SEQUENCE 9 AA: 1038 MW: 69FE5B9C75BB1B CRC64:

Query Match 26.7%; Score 16; DB 1; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.1e+05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 FSG 12
 DB 6 FSG 8

RESULT 10
 URAL_HUMAN
 ID URAL_HUMAN STANDARD: PRT; 10 AA.
 AC P32118;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Unknown protein from 2D-page of red blood cells (Spot 1) (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.
 RC TISSUE-Erythrocyte;
 RX MEDLINE=94147970; PubMed=8313871;
 RA Golar O., Hughes G.J., Frutiger S., Paquet N., Balroch A.,
 RA Pasquail C., Sanchez J.-C., Tissot J.-D., Appel R.D., Walzer C.,
 RA Balant L., Hochstrasser D.F.,
 RL "Plasma and red blood cell protein maps: update 1993."
 RL Electrophoresis 14:1223-1231(1993).
 CC -1 MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
 CC PROTEIN IS: 5.4, ITS MW IS: 23 Kda.
 DR SWISS-2DPAGE: P32118; HUMAN.
 DR NON_TER 1 1
 FT NON_TER 1 1
 SO SEQUENCE 10 AA: 977 MW: 723G65B1ADD0587B CRC64:

Query Match 26.7%; Score 16; DB 1; Length 10;
 Best Local Similarity 60.0%; Pred. No. 2.4e+03;
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 3 EGI DT 7
 DB 3 EGIAT 7

RESULT 11
 TKNA_CHICK
 ID TKNA_CHICK STANDARD: PRT; 11 AA.
 AC P19850;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Substance P.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorphi; Paracanthopterygii; Gadiformes; Gadidae; Gadus.
 OX NCBI_TaxID=8049;
 RN [1]
 RP SEQUENCE.
 RC TISSUE-Brain;
 RX MEDLINE=92288992; PubMed=1376687;
 RA Jensen J., Conlon J.M.,
 RT "Substance-P-related and neurokinin-A-related peptides from the brain
 RT of the cod and trout."
 RL Eur. J. Biochem. 206:659-664(1992).
 CC -1 FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
 CC EVOLVE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
 CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
 CC MUSCLES.
 CC -1 SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
 DR InterPro: IPR003580; Protachykinin.
 DR InterPro: IPR002040; Tachykinin.
 DR Pfam: PF02202; Tachykinin; 1.
 DR SMART: SM00203; TK; 1.
 DR PROSITE: PS00267; TACHYKININ; 1.
 KW Tachykinin; Neuropeptide; Amidation; Neurotransmitter.
 FT MOD_RES 11 11
 SO SEQUENCE 11 AA: 1315 MW: 214860D75996C6C7 CRC64:

Query Match 26.7%; Score 16; DB 1; Length 11;
 Best Local Similarity 60.0%; Pred. No. 2.7e+03;
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 8 QRFSG 12

CC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE.
 RX TISSUE-Intestine; PubMed=2452461;
 RX MEDLINE=88204263; PubMed=2452461;
 RA Conlon J.M., Katsoulis S., Schmidt W.E., Thim L.,
 RT "A[ar3] substance P and neurokinin A from chicken small intestine."
 RL Regul. Pept. 20:171-180(1988).
 CC -1 FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
 CC EVOLVE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
 CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
 CC MUSCLES.
 CC -1 SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
 DR PIR: JN0023; JN0023.
 DR InterPro: IPR003580; Protachykinin.
 DR InterPro: IPR002040; Tachykinin.
 DR Pfam: PF02202; Tachykinin; 1.
 DR SMART: SM00203; TK; 1.
 DR PROSITE: PS00267; TACHYKININ; 1.
 KW Tachykinin; Neuropeptide; Amidation; Neurotransmitter.
 FT MOD_RES 11 11
 SO SEQUENCE 11 AA: 1377 MW: 21487FE3C9D6C6C7 CRC64:

Query Match 26.7%; Score 16; DB 1; Length 11;
 Best Local Similarity 60.0%; Pred. No. 2.7e+03;
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 8 QRFSG 12

RESULT 12
 TKNA_GADMO
 ID TKNA_GADMO STANDARD: PRT; 11 AA.
 AC P28498;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Substance P.
 OS Gadus morhua (Atlantic cod).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorphi; Paracanthopterygii; Gadiformes; Gadidae; Gadus.
 OX NCBI_TaxID=8049;
 RN [1]
 RP SEQUENCE.
 RC TISSUE-Brain;
 RX MEDLINE=92288992; PubMed=1376687;
 RA Jensen J., Conlon J.M.,
 RT "Substance-P-related and neurokinin-A-related peptides from the brain
 RT of the cod and trout."
 RL Eur. J. Biochem. 206:659-664(1992).
 CC -1 FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
 CC EVOLVE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
 CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
 CC MUSCLES.
 CC -1 SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
 DR InterPro: IPR003580; Protachykinin.
 DR InterPro: IPR002040; Tachykinin.
 DR Pfam: PF02202; Tachykinin; 1.
 DR SMART: SM00203; TK; 1.
 DR PROSITE: PS00267; TACHYKININ; 1.
 KW Tachykinin; Neuropeptide; Amidation; Neurotransmitter.
 FT MOD_RES 11 11
 SO SEQUENCE 11 AA: 1315 MW: 214860D75996C6C7 CRC64:

Query Match 26.7%; Score 16; DB 1; Length 11;
 Best Local Similarity 60.0%; Pred. No. 2.7e+03;
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 8 QRFSG 12

```

Db          5 QQFCG 9

RESULT 13
TKNA_HORSE
ID          STANDARD:      PRT:      11 AA.
AC          P01290:
DT          21-JUL-1986 (Rel. 01, Created)
DT          21-JUL-1986 (Rel. 01, Last sequence update)
DT          30-MAY-2000 (Rel. 39, Last annotation update)
DE          Substance P.
GN          TAC1 OR NKNA OR TAC2 OR NKA.
OS          Equus caballus (horse), and
OC          Cavaria porcellus (Guinea pig).
OC          Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC          Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX          NCBI_TaxId=9796, 10141;
RN          [1]
RP          SEQUENCE.
RC          SPECIES=Horse;
RA          Studer R.O., Trzeciak A., Lergier W.;
RT          "Isolation and amino-acid sequence of substance P from horse
RT          intestine.";
RL          Helv. Chim. Acta 56:860-866(1973).
RN          [2]
RP          SEQUENCE.
RC          SPECIES=C.porcellus;
RX          MEDLINE=90044685; PubMed=2478925;
RA          Murphy R.;
RT          "Primary amino acid sequence of guinea-pig substance P.";
RL          Neuropeptides 14:105-110(1989).
CC          -1- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
CC          EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
CC          SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
CC          MUSCLES.
CC          -1- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
DR          PIR: A01558; SPHO.
DR          PIR: A0654; A0654.
DR          InterPro: IPR003580; Protachykinin.
DR          InterPro: IPR002040; Tachykinin.
DR          Pfam: PF02202; Tachykinin; 1.
DR          SMART: SM00203; TK; 1.
DR          PROSITE: PS00267; TACHYKININ; 1.
KM          Tachykinin; Neuropeptide; Neurotransmitter.
FT          MOD_RES      11
FT          MOD_RES      11 AMIDATION.
SQ          SEQUENCE      11 AA; 1349 MW; 3E757EE3C9DC6C7 CRC64;

Query Match
Best Local Similarity      26.7%; Score 16; DB 1; Length 11;
Matches      3; Conservative      1; Mismatches      1; Indels      0; Gaps      0;

OY          8 ORFSG 12
Db          5 QQFCG 9

RESULT 14
ALL5_CYPDO
ID          STANDARD:      PRT:      8 AA.
AC          P82156:
DT          30-MAY-2000 (Rel. 39, Created)
DT          30-MAY-2000 (Rel. 39, Last sequence update)
DT          30-MAY-2000 (Rel. 39, Last annotation update)
DE          Cydiastatin 5.
OS          Cydia pomonella (Codling moth).
OC          Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC          Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata;
OC          Ditrysia; Tortricidae; Tortricidae; Olethreutinae; Cydia.
OX          NCBI_TaxId=82600;
RN          [1]
RP          SEQUENCE.
RC          TISSUE=Larva;

```

```

RX          MEDLINE=98054539; PubMed=9392829;
RA          Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Winstanley D.,
RA          Davey M., East P.D., Thorpe A.;
RT          "Lepidopteran peptides of the allatostatin superfamily.";
RL          Peptides 18:1301-1309(1997).
CC          -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KM          Neuropeptide; Amidation.
FT          MOD_RES      8
FT          MOD_RES      8 AMIDATION.
SQ          SEQUENCE      8 AA; 898 MW; 922879CABB58640D CRC64;

Query Match
Best Local Similarity      25.0%; Score 15; DB 1; Length 8;
Matches      3; Conservative      0; Mismatches      2; Indels      0; Gaps      0;

OY          2 AEGUD 6
Db          1 ARGVD 5

RESULT 15
FAR8_CALVO
ID          STANDARD:      PRT:      8 AA.
AC          P41863;
DT          01-NOV-1995 (Rel. 32, Created)
DT          01-NOV-1995 (Rel. 32, Last sequence update)
DT          01-NOV-1995 (Rel. 32, Last annotation update)
DE          Callifmrfamide 8.
OS          Calliphora vomitoria (Blue blowfly).
OC          Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC          Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC          Muscomorpha; Oestroidea; Calliphoridae; Calliphora.
OX          NCBI_TaxId=27454;
RN          [1]
RP          SEQUENCE.
RC          TISSUE=Thoracic ganglion;
RX          MEDLINE=92196111; PubMed=1549595;
RA          Duve H., Johnsen A.H., Sewell J.C., Scott A.G., Orchard I.,
RA          Rehfeld J.F., Thorpe A.;
RT          "Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH2
RT          neuropeptides (designated callifmrfamides) from the blowfly
RT          Calliphora vomitoria.";
RL          Proc. Natl. Acad. Sci. U.S.A. 89:2326-2330(1992).
CC          -1- SIMILARITY: BELONGS TO THE FAR8 (FMRFAMIDE RELATED PEPTIDE)
CC          FAMILY.
DR          PIR: H41978; H41978.
KM          Neuropeptide; Amidation.
FT          MOD_RES      8
FT          MOD_RES      8 AMIDATION.
SQ          SEQUENCE      8 AA; 957 MW; 72D40699CAA44D8 CRC64;

Query Match
Best Local Similarity      25.0%; Score 15; DB 1; Length 8;
Matches      3; Conservative      0; Mismatches      2; Indels      0; Gaps      0;

OY          6 DFOGF 10
Db          4 DFWRF 8

Search completed: April 15, 2003, 10:20:22
Job time : 12 secs

```


GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 15, 2003, 10:17:20 ; Search time 28 Seconds

(without alignments)
88.306 Million cell updates/sec

Title: US-09-521-527c-1

Perfect score: 60

Sequence: 1 AAEGDTRFSG 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 segs, 206047115 residues

Total number of hits satisfying chosen parameters: 2041

Minimum DB seq length: 0
Maximum DB seq length: 12

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL.21:*
1: SP_Archea:*
2: SP_Bacteria:*
3: SP_Fungi:*
4: SP_Human:*
5: SP_Invertebrate:*
6: SP_Mammal:*
7: SP_Mhc:*
8: SP_Organelle:*
9: SP_Phage:*
10: SP_Plant:*
11: SP_Rodent:*
12: SP_Virus:*
13: SP_Vertebrate:*
14: SP_Unclassified:*
15: SP_RVirus:*
16: SP_Bacteriap:*
17: SP_Archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	36.7	11	5	Q9TWM2
2	21	35.0	12	12	Q86570
3	19	31.7	8	5	Q94695
4	19	31.7	10	12	P90373
5	19	31.7	11	2	Q8RKN1
6	18	30.0	8	2	Q9R3X0
7	18	30.0	9	15	Q70140
8	18	30.0	10	11	Q9QVE6
9	18	30.0	11	9	Q9TOR6
10	17	28.3	8	4	Q9Y4J4
11	17	28.3	8	7	Q2S801
12	17	28.3	8	7	Q2S797
13	17	28.3	8	7	Q2S812
14	17	28.3	11	10	Q04131
15	17	28.3	11	10	Q9H3W2
16	17	28.3	12	6	Q9TRV6
17	17	28.3	12	6	Q9TRV6

17	16	26.7	8	4	Q15895	Q15895 homo sapien
18	16	26.7	9	4	Q96P97	Q96P97 homo sapien
19	16	26.7	9	12	Q88612	Q88612 middelburg
20	16	26.7	10	2	Q8VN85	Q8VN85 helicobacte
21	16	26.7	11	5	Q23875	Q23875 dictyostell
22	16	26.7	11	6	P83128	P83128 bos indicus
23	16	26.7	11	6	Q9TRX0	Q9TRX0 sus scrofa
24	16	26.7	11	7	Q77899	Q77899 oreochromis
25	16	26.7	11	7	Q77900	Q77900 oreochromis
26	16	26.7	11	7	Q77901	Q77901 oreochromis
27	16	26.7	11	7	Q77902	Q77902 oreochromis
28	16	26.7	11	7	Q77903	Q77903 oreochromis
29	16	26.7	11	7	Q77904	Q77904 oreochromis
30	16	26.7	11	7	Q77905	Q77905 oreochromis
31	16	26.7	11	7	Q77916	Q77916 oreochromis
32	16	26.7	11	7	Q77917	Q77917 oreochromis
33	16	26.7	11	7	Q77921	Q77921 pseudotroph
34	16	26.7	11	16	Q9K7A4	Q9K7A4 bacillus ha
35	16	26.7	12	2	Q93U04	Q93U04 escherichia
36	16	26.7	12	10	P82328	P82328 pisum sativ
37	16	26.7	12	11	Q925V7	Q925V7 mus musculu
38	15	25.0	8	5	Q94623	Q94623 manduca sex
39	15	25.0	9	2	Q43960	Q43960 azotobacter
40	15	25.0	9	2	P72149	P72149 pseudomonas
41	15	25.0	9	10	Q9AXH8	Q9AXH8 mesembryant
42	15	25.0	10	2	Q9L5W6	Q9L5W6 liberibacte
43	15	25.0	10	2	Q9S3J6	Q9S3J6 escherichia
44	15	25.0	10	2	Q9R5N1	Q9R5N1 clostridium
45	15	25.0	10	2	Q9R5N1	Q9R5N1 clostridium

ALIGNMENTS

RESULT 1	Q9TWM2	PRELIMINARY;	PRT;	11 AA.
ID	Q9TWM2			
AC	Q9TWM2;			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-MAY-2000 (TREMBLrel. 13, Last annotation update)			
DE	Buccalin B, BUCB.			
OS	Aplysia californica (California sea hare).			
OC	Eukaryota; Metazoa; Mollusca; Gastropoda; Opisthobranchia; Anaspidia;			
OC	Aplysiidae; Aplysia.			
OX	NCBI_Taxid=6500;			
RN	[1]			
RP	SEQUENCE.			
RX	MEDLINE=95083478; PubMed=7991459;			
RX	Vilim F.S., Cropper E.C., Rosen S.C., Tenenbaum R., Kupfermann I.,			
RA	Weiss K.R.;			
RT	"Structure, localization, and action of buccalin B: a bioactive			
RT	peptide from Aplysia."			
RL	Peptides 15:959-969(1994).			
SQ	SEQUENCE 11 AA; 1153 MW; 692253F9C9C86B44 CRC64;			
Query Match	36.7%;	Score 22;	DB 5;	Length 11;
Best local Similarity	55.6%;	Pred. No. 1.2e+03;		
Matches	5;	Conservative	0;	Mismatches 4;
				Indels 0;
				Gaps 0;
QY	4 GLDTRFSG 12			
DB	1 GLDTRFSG 9			
RESULT 2	Q86570	PRELIMINARY;	PRT;	12 AA.
ID	Q86570;			
AC	Q86570;			
DT	01-NOV-1996 (TREMBLrel. 01, Created)			
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)			
DT	01-NOV-1998 (TREMBLrel. 08, Last annotation update)			
DE	Small hepatitis delta antigen (Fragment).			

```

OS Hepatitis delta virus (HDV).
OC Viruses; Deltavirus.
OX NCBI_TaxID=12475;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=93190623; Pubmed=8447179;
RX Khudvyakov Yu.E., Favovov M.O., Fields H.A.;
RT "A small open reading frame of the hepatitis delta virus antigenomic
RT RNA encodes a protein that elicits antibodies in some infected
RT patients.";
RL Virus Res. 27:13-24(1993).
DR EMBL; S56617; AAB25696.1; -.
FT NON_TER 1
SQ SEQUENCE 12 AA; 1330 MW; 0E63E3A15486400 CRC64;

Query Match 35.0%; Score 21; DB 12; Length 12;
Best Local Similarity 100.0%; Pred. No. 2e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 EGLD 6
Db 1 EGLD 4

RESULT 3
Q94695 PRELIMINARY; PRT; 8 AA.
ID Q94695;
AC Q94695;
DT 01-FEB-1997 (TReMBLrel. 02, Created)
DT 01-FEB-1997 (TReMBLrel. 02, Last sequence update)
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
DE Actin (Fragment).
GN ARDC.
OS Physarum polycephalum (Slime mold).
OC Eukaryote; Mycetozoa; Myxogastria; Myxogastromycetidae; Physarida;
OC Physarum.
OX NCBI_TaxID=5791;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96182101; Pubmed=8622700;
RA Bernard M., Lagneel C., Pallotta D., Pierron G.;
RT "Mapping of a replication origin within the promoter region of two
RT unlinked, abundantly transcribed actin genes of Physarum
RT polycephalum.";
RL Mol. Cell. Biol. 16:968-976(1996).
DR EMBL; M73459; AAB03706.1; -.
FT NON_TER 8
SQ SEQUENCE 8 AA; 878 MW; F4C6C2CAAB187B16 CRC64;

Query Match 31.7%; Score 19; DB 5; Length 8;
Best Local Similarity 66.7%; Pred. No. 6.7e+05;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 EGLDTQ 8
Db 2 EGEDVQ 7

RESULT 4
P90373 PRELIMINARY; PRT; 10 AA.
ID P90373;
AC P90373;
DT 01-MAY-1997 (TReMBLrel. 03, Created)
DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE U5 (Fragment).
OS Pseudorabies virus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10345;
RN [1]
RP SEQUENCE FROM N.A.
DR STRAIN=BECKER;

```

```

RA Pederson N.E., Casey J.T., II, Koslowski K.M.;
RT "Pseudorabies virus UL6 and UL7 sequence.";
RL Submitted (AUG-1996) to the EMBL/Genbank/DBJ databases.
DR EMBL; U66829; AAB37266.1; -.
FT NON_TER 10
SQ SEQUENCE 10 AA; 859 MW; D232421DDDDDDDD CRC64;

Query Match 31.7%; Score 19; DB 12; Length 10;
Best Local Similarity 100.0%; Pred. No. 4e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAE6 4
Db 6 AAE6 9

RESULT 5
Q8RKN1 PRELIMINARY; PRT; 11 AA.
ID Q8RKN1;
AC Q8RKN1;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Beta-lactamase CTX-M-9 (Fragment).
GN BLACTX-M-9.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=743-D;
RA Sabate M., Navarro F., Miro E., Campoy S., Mirelis B., Barbe J.,
RA Prats G.;
RT "A novel complex sul1-type integron in Escherichia coli carrying the
RT bla(CTX-M-9) gene.";
RL Submitted (MAR-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL; AI092058; AAM15718.1; -.
FT NON_TER 1
SQ SEQUENCE 11 AA; 1071 MW; C26BF418D050440D CRC64;

Query Match 31.7%; Score 19; DB 2; Length 11;
Best Local Similarity 100.0%; Pred. No. 4.5e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AEGL 5
Db 8 AEGL 11

RESULT 6
Q9R3X0 PRELIMINARY; PRT; 8 AA.
ID Q9R3X0;
AC Q9R3X0;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE Ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit
DE (Fragment).
GN RECL.
OS Bacteria; Cyanobacteria; Oscillatoriales; Planctothrix.
OC Bacteria; Cyanobacteria; Oscillatoriales; Planctothrix.
OX NCBI_TaxID=59512;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BC-PLA 9316, AND BC-PLA 9303;
RX MEDLINE=20005589; Pubmed=10537197;
RA Beard S.J., Handley B.A., Hayes P.K., Walsby A.E.;
RT "The diversity of gas vesicle genes in Planctothrix rubescens from
RT Lake Zurich.";
RL Microbiology 145:2757-2768(1999).
DR EMBL; AJ132249; CAB59537.1; -.
DR EMBL; AJ132248; CAB59534.1; -.

```



```

FT NON_TER 1 1
SQ SEQUENCE 8 AA; 957 MW; 33DIAAA685HB19CB CRC64;

Query Match
Best Local Similarity 30.0%; Score 18; DB 2; Length 8;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 EGLDPT 7
   1 11
   1 11
Db 3 ESMdT 7

RESULT 7
ID 070140 PRELIMINARY; PRT; 9 AA.
AC 070140:
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Tat protein (Fragment).
GN TAT.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=95194694; PubMed=7888189;
RA Gao F., Yue L., Craig S., Thornton C.L., Robertson D.L.,
RA McCutchan F.E., Bradac J.A., Sharp P.M., Hahn B.H.;
RT "Genetic variation of HIV type 1 in four World Health Organization-
RT sponsored vaccine evaluation sites: generation of functional envelope
RT (glycoprotein 160) clones representative of sequence subtypes A, B, C,
RT and E. WHO Network for HIV Isolation and Characterization.";
RL AIDS Res. Hum. Retroviruses 10:1359-1368(1994).
RN 12
RP SEQUENCE FROM N.A.
RX MEDLINE=96190564; PubMed=8627686;
RA Gao F., Morrison S.G., Robertson D.L., Thornton C.L., Craig S.,
RA Karlsson G., Sodroski J., Morgado M., Galvao-Castro B.,
RA von Briesen H., Beddows S., Weber J., Sharp P.M., Shaw G.M.,
RA Hahn B.H.;
RT "Molecular cloning and analysis of functional envelope genes from
RT human immunodeficiency virus type 1 sequence subtypes A through G. The
RT WHO and NIAID Networks for HIV Isolation and Characterization.";
RL J. Virol. 70:1651-1657(1996).
RN 13
RP SEQUENCE FROM N.A.
RX STRAIN=020;
RA Allen E.E.;
RL Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL: U08794; AAB05175.1; -.
FT NON_TER 1 1
SQ SEQUENCE 9 AA; 1098 MW; 5B76D40AB1AB01A3 CRC64;

Query Match
Best Local Similarity 30.0%; Score 18; DB 15; Length 9;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 DTORF 10
   1 11
   1 11
Db 4 ETDRF 8

RESULT 8
ID 090VE6 PRELIMINARY; PRT; 10 AA.
AC 090VE6:
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Protamine MP2 Intermediate protein PMP2/26 (Fragment).

```

```

OS Mus sp.
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10095;
RN 11
RP SEQUENCE.
RX MEDLINE=92174934; PubMed=1541289;
RA Chauviere M., Martinge A., Debarle M., Sautiere P., Chevallier P.;
RT "Molecular characterization of six intermediate proteins in the
RT processing of mouse protamine P2 precursor.";
RL Eur. J. Biochem. 204:759-765(1992).
FT NON_TER 1 1
FT NON_TER 10 10
SQ SEQUENCE 10 AA; 1028 MW; 2B099C75B72866D8 CRC64;

Query Match
Best Local Similarity 30.0%; Score 18; DB 11; Length 10;
Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 EGLDPTOR 9
   1 11
   1 11
Db 4 QGLSPER 10

RESULT 9
ID 09TOR6 PRELIMINARY; PRT; 11 AA.
AC 09TOR6:
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2002 (TReMBLrel. 20, Last sequence update)
DT 01-MAY-2002 (TReMBLrel. 20, Last annotation update)
DE Ner protein (Fragment).
OS Bacteriophage Mu.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
OC Mu-like viruses.
OX NCBI_TaxID=10677;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=85153064; PubMed=6099259;
RA Krause H.M., Higgins N.P.;
RT "On the Mu repressor and early DNA intermediates of transposition.";
RL Cold Spring Hard. Symp. Quant. Biol. 49:827-834(1984).
RN 12
RP SEQUENCE FROM N.A.
RX MEDLINE=92356834; PubMed=1386645;
RA Alazard R., Beliermier M., Chandelier M.;
RT "Escherichia coli integration host factor stabilizes bacteriophage Mu
RT repressor interactions with operator DNA in vitro.";
RL Mol. Microbiol. 6:1707-1714(1992).
RN 13
RP SEQUENCE FROM N.A.
RX MEDLINE=92356835; PubMed=1386646;
RA Gama M.J., Tousseint A., Higgins N.P.;
RT "Stabilization of bacteriophage Mu repressor-operator complexes by the
RT Escherichia coli integration host factor protein.";
RL Mol. Microbiol. 6:1715-1722(1992).
DR EMBL: M10192; AAA32371.2; -.
FT NON_TER 11 11
SQ SEQUENCE 11 AA; 1006 MW; 89978F816766987A CRC64;

Query Match
Best Local Similarity 30.0%; Score 18; DB 9; Length 11;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAEGL 5
   1 11
   1 11
Db 2 AADGM 6

RESULT 10
ID 09Y4J4 PRELIMINARY; PRT; 8 AA.
AC 09Y4J4:

```

DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
 DE Runt/82nt/MTG8 protein (Fragment).
 GN Runt/82nt/MTG8.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9502916; PubMed=7919324;
 RA Tjorne J.E., Calabi F.;
 RT "Alternative, out-of-frame runt/MTG8 transcripts are encoded by the
 RT derivative (8) chromosome in the t(8;21) of acute myeloid leukemia
 RT M2.";
 RL Blood 84:2115-2121(1994).
 DR EMBL: S74092; AAD14144.1; -.
 FT NON-TER 1 1
 SQ SEQUENCE 8 AA; 1067 MW; 20F414044B17244B CRC64;

Query Match 28.3%; Score 17; DB 7; Length 8;
 Best Local Similarity 42.9%; Pred. No. 6.7e+05;
 Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 3 EGIJTOR 9
 DB 1 ENLENNR 7

RESULT 11
 ID Q29801 PRELIMINARY; PRT; 8 AA.

AC Q29801;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE HLA class II DR-beta chain (Fragment).
 GN HLA-DRB1*15.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96175156; PubMed=9110934;
 RA Svensson A.C., Setterblad N., Pihlgren U., Rask L., Andersson G.;
 RT "Evolutionary relationship between different human major
 RT histocompatibility complex HLA-DR haplotypes.";
 RL Immunogenetics 43:304-314(1996).
 DR EMBL: X88791; CAA61270.1; -.
 FT NON-TER 1 1
 SQ SEQUENCE 8 AA; 953 MW; C2C6C5B9CA731AA CRC64;

Query Match 28.3%; Score 17; DB 7; Length 8;
 Best Local Similarity 66.7%; Pred. No. 6.7e+05;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 6 DTORFS 11
 DB 1 DTLDPS 6

RESULT 12
 ID Q29797 PRELIMINARY; PRT; 8 AA.
 AC Q29797;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE HLA class II DR-beta chain (Fragment).
 GN HLA-DRB1*01.

OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96175156; PubMed=9110934;
 RA Svensson A.C., Setterblad N., Pihlgren U., Rask L., Andersson G.;
 RT "Evolutionary relationship between different human major
 RT histocompatibility complex HLA-DR haplotypes.";
 RL Immunogenetics 43:304-314(1996).
 DR EMBL: X88793; CAA61272.1; -.
 FT NON-TER 1 1
 SQ SEQUENCE 8 AA; 953 MW; C2C6C5B9CA731AA CRC64;

Query Match 28.3%; Score 17; DB 7; Length 8;
 Best Local Similarity 66.7%; Pred. No. 6.7e+05;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 6 DTORFS 11
 DB 1 DTLDPS 6

RESULT 13
 ID Q29812 PRELIMINARY; PRT; 8 AA.
 AC Q29812;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE HLA class II DR-beta chain (Fragment).
 GN HLA-DRB6*01.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96175156; PubMed=9110934;
 RA Svensson A.C., Setterblad N., Pihlgren U., Rask L., Andersson G.;
 RT "Evolutionary relationship between different human major
 RT histocompatibility complex HLA-DR haplotypes.";
 RL Immunogenetics 43:304-314(1996).
 DR EMBL: X88794; CAA61273.1; -.
 FT NON-TER 1 1
 SQ SEQUENCE 8 AA; 953 MW; C2C6C5B9CA731AA CRC64;

Query Match 28.3%; Score 17; DB 7; Length 8;
 Best Local Similarity 66.7%; Pred. No. 6.7e+05;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 6 DTORFS 11
 DB 1 DTLDPS 6

RESULT 14
 ID Q04131 PRELIMINARY; PRT; 11 AA.
 AC Q04131;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Wound induced protein (Tomato).
 OS Lycopersicon esculentum (Tomato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 OX NCBI_TaxID=4081;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=PIK-RED; TISSUE=PERICARP;
 RX MEDLINE=91355936; PubMed-1715787;
 RA Parsons B.L., Mattoo A.K.;
 RT "Wound regulated accumulation of specific transcripts in tomato
 RT fruit: interactions with fruit development, ethylene and light.";
 RL Plant Mol. Biol. 17:453-464(1991).
 DR EMBL; X59884; CAA42539.1; -.
 FT NON_TER
 SQ SEQUENCE 11 AA; 1278 MW; 92CB25782873325 CRC64;

Query Match 28.3%; Score 17; DB 10; Length 11;
 Best Local Similarity 75.0%; Pred. No. 1.1e+04;
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 EGLD 6
 :|||
 Db 6 KGLD 9

RESULT 15

O9H3W2 PRELIMINARY; PRT; 12 AA.
 AC O9H3W2;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE TRE17 protein.
 GN TRE17 GENE.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94010310; PubMed-8406013;
 RA Onno M., Nakamura T., Hillova J., Hill M.;
 RT "Identification of novel sequences in the repertoire of hypervariable
 RT TRE17 genes from immortalized nonmalignant and malignant human
 RT keratinocytes.";
 RL Gene 131:209-215(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93228825; PubMed-8471161;
 RA Onno M., Nakamura T., Marriage-Samson R., Hillova J., Hill M.;
 RT "Human TRE17 oncogene is generated from a family of homologous
 RT polymorphic sequences by single-base changes.";
 RL DNA Cell Biol. 12:107-118(1993).
 DR EMBL; X71371; CAC16150.1; -.
 SQ SEQUENCE 12 AA; 1252 MW; 6E17AFDE83ADD87B CRC64;

Query Match 28.3%; Score 17; DB 4; Length 12;
 Best Local Similarity 60.0%; Pred. No. 1.2e+04;
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 3 EGLDT 7
 :|||
 Db 5 EGADS 9

Search completed: April 15, 2003, 10:20:58
 Job time : 30 secs

